Benefit substitution and heterogenous effects of stricter criteria for disability insurance

Tunga Kantarcı*, Jan-Maarten van Sonsbeek†, and Yi Zhang‡

9 April 2021

Abstract

We evaluate a major disability insurance (DI) reform introduced in 2006 in the Netherlands which tightened eligibility criteria and introduced (financial) incentives for work resumption for both the employees and employers. Based on administrative data on the universe of individuals who fall sick before and after the reform, difference-in-difference regressions show that the reform reduces DI receipt substantially and increases labor participation and unemployment insurance (UI) receipt to a limited extent. It increases wages earned, which fully compensate lost DI benefits, and to a lesser extent also the amount of UI received. The effects vary substantially with the individuals' age, work status, type of employment contract and occupation. The impact of the reform is persistent over time for DI, labor participation and earnings.

1 Introduction

As the number of DI recipients grow in many western countries, evidence on the impact of DI reforms on benefit claiming and work resumption are of utmost importance to policy makers. Figure 1 shows the number of people receiving benefits in DI in ten OECD countries during the period from 2007 to 2016. While many of the countries show an increasing trend over the observation period, in a small number of countries inflow into the disability scheme has decreased due to the major policy reforms implemented in these countries during the last two decades. Furthermore, during the period from 2007 to 2015, public spending on incapacity accounted, on average, for 2.06 percent of the gross domestic product in the OECD countries (OECD, 2018).

Through disability reforms, governments introduced different types of measures to reduce DI claiming and increase labor participation among sick individuals. Existing studies focus on analyzing the effects of two main measures. The first is tightening the eligibility criteria (Autor and Duggan, 2003; Karlström et al., 2008; De Jong et al., 2011; Staubli, 2011; Campolieti and Riddell, 2012; Borghans et al., 2014; Moore, 2015; Autor et al., 2016; Hullegie and Koning, 2018), and the second is reducing the benefit level (Gruber, 2000; Campolieti, 2004; Marie and Vall

*Department of Econometrics and Operations Research, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands, and Netspar (e-mail: kantarci@tilburguniversity.edu)

†Department of Public Finance, Netherlands Bureau for Economic Policy Analysis, P.O. Box 80510, 2508 GM The Hague, The Netherlands, and Netspar (e-mail: j.m.van.sonsbeek@cpb.nl)

‡China Center for Human Capital and Labor Market Research, Central University of Finance and Economics, 39 South College Road, Haidian District, 100081, Beijing, China (e-mail: yi.zhang@cufe.edu.cn)

Castello, 2012; Kostøl and Mogstad, 2014; Low and Pistaferri, 2015; Deshpande, 2016; Mullen and Staubli, 2016; Fevang et al., 2017; Koning and van Sonsbeek, 2017; Zaresani, 2018; Deuchert and Eugster, 2019; Ruh and Staubli, 2019). The main finding in this literature is that restricting entitlement and reducing the benefit level reduce DI receipt and increase labor participation. Some also find that reforms induce spillover effects into alternative benefit programs.[1]

In the early 2000s the Netherlands had become one of the countries with the highest fraction of disabled workers in the working population, as the total number of DI recipients reached almost one million whereas the working population was around 7 million. Successive governments implemented a series of radical reforms to limit the access to DI. In 2006, the Work and Income According to Labor Capacity Act (WIA) came into effect as the final element of these reforms, as a successor to the Disability Insurance Act (WAO). Foundations of the WIA were laid in 2001 when a committee of experts from main political parties advised the then government on DI reform. Labour participation increase was framed as the prime goal of the new law, whereas generating budgetary savings by reducing DI use came second. This order followed a broader trend in OECD countries where between 1985 and 2000 DI policies shifted their focus from compensation to integration (OECD, 2003). The WIA extended the sickness insurance (SI) scheme that precedes the DI scheme from one to two years, introduced stricter criteria to enter the DI scheme, but also introduced (financial) incentives for work resumption in the DI scheme for both the employees and employers. Extension of the sickness period makes the WIA reform distinctive compared to disability reforms in other countries. The strong incentive for the employer to facilitate work resumption is that it is obliged to compensate the employee for wage loss during the two-year period of the sickness scheme. However, mandating employers to pay wages for a longer period may also make them reluctant to hire workers with health problems (Koning and Lindeboom, 2015). These employer incentives during the sickness period, however, precede the incentives built into the disability scheme, and therefore could affect individual behavior during the disability period. Overall, the WIA is a comprehensive DI reform that introduced unique incentives for work resumption that are seldom seen in disability reforms in other countries. It is studied by other countries facing unsustainable growth in their DI schemes, such as the US (Burkhauser et al., 2014; Fultz, 2015).

Up till now, research analyzing the impact of the WIA reform is limited to Van Sonsbeek and Gradus (2013).[2] Based on aggregated quarterly data on DI use, they find that the whole of the WIA reform, including the extension of the sickness period besides the other incentives it introduced for work resumption, has led to a sharp fall in the number of new benefit receipts, in addition to what has already been achieved through previous reforms, generating large bud-

[1] Instead of exploiting exogenous variation in benefit rules due to policy reforms, Chen and van der Klaauw (2008), Maestas et al. (2013), French and Song (2014), and Gelber et al. (2017) exploit variation in benefit eligibility rules or benefit levels imbed in DI programs to analyze the causal effect of benefit incentives on benefit receipt and labor supply. A smaller literature investigates how incentive changes on the employer side influence DI receipt. For example, De Jong and Lindeboom (2004) show that mandating firms to use preventive and reintegration measures to reduce sickness absenteeism does not decrease absence rates. Koning (2009) and De Groot and Koning (2016) analyze the introduction and abolishment of experience rating for firms' DI premium, and find that experience rating effectively decreases DI receipt. Another strand of the literature pays attention to the non-economic outcomes of disability reforms (Dahl and Gielen, 2018; García-Gómez and Gielen, 2018). García-Gómez and Gielen, for example, find that despite the gains in public finances, stricter eligibility criteria reduce life expectancy among women.

[2] The lack of research on the impact of the WIA is due to the lack of data on sick individuals who subsequently could file disability applications. The sickness benefit was reformed in 1994, 1996 and 2004 to mandate the employer to pay during the sickness period 70 percent of the earnings before sickness. Since no sickness benefit is paid by the government, but wage is paid by the employer, there is no registration of sickness absence by the government since these reforms. New reintegration regulations for employers were introduced in the sickness scheme in 2002 ("Gatekeeper protocol"). Only after this year the government started to register sickness cases to monitor whether employers comply with the new regulations.

getary savings for the government. In their ex-ante evaluation of the reform, Van Sonsbeek and Gradus predicted that on average incomes of DI claimants do not decrease because the gains from additional labor participation compensate for the loss of DI benefits.

Some of the main questions from the policy debate, and the claims from the ex-ante evaluation of the reform, remain to be examined. First, although it is known that DI receipt strongly decreased after the reform, it is not analyzed to which extent the decrease in DI use among sick individuals led to an increase in labour participation or use of benefits from alternative benefit programs such as the unemployment benefit. Second, it is not clear how the amounts of earnings and benefits received by sick-listed workers are affected by the reform. Although it was claimed at the onset of the reform that the income position of sick-listed workers would remain stable, it was never confirmed that these workers are able to compensate for loss of DI benefits by other means. Third, it is not known whether the effects of the WIA reform are structural or fade in the long run, for example because people who do not first enter the WIA later become more sick and still become incapacitated for work. It is known that DI receipt had risen since the onset of the reform (Berendsen et al., 2019), but this need not be caused by diminishing reform effects over time, but it can be caused, for example, by increasing labor participation of older workers who have a higher DI risk. Finally, it is not known how the effects of the WIA reform vary across subgroups of sick individuals. In the Netherlands, this is of particular interest because, like many other continental European countries, the Dutch labor market is characterized by a strong insider-outsider segmentation (Hausermann and Schwander, 2012) and low labor mobility of older workers (Visser et al., 2018).

In this study we exploit unique administrative data from the Employee Insurance Agency (UWV) on the universe of individuals who fell sick in the third and fourth quarters of 2003 and the first quarter of 2004. The two groups of individuals who fell sick in the last two quarters of 2003 are insured under the old WAO scheme and subject to different rules, while the third group of individuals who fell sick in the beginning of 2004 are insured under the WIA scheme and subject to additional and new rules. To investigate the effects of the stricter criteria of the WIA regime, we compare the work and benefit claiming behavior of the three groups of individuals before they fall sick and after they become eligible to apply for DI.[3] The three groups of sick individuals are comparable in background characteristics, and economic shocks are likely to affect the behavior of the three groups in similar ways since eligibility for different disability schemes is determined by falling sick within close proximity in time. This allows to attribute the differences in the work and benefit claiming behavior across the three groups to the differences in the rules of the disability schemes that apply to these groups. We take a difference-in-difference approach to identify the causal effects of the disability reform.

A growing literature relies on DI reforms to provide causal evidence on the impact of DI policies. Karlström et al. (2008) study the effect of abolishing generous eligibility rules for DI in the Swedish DI program in 1997. Staubli (2011) studies the effect of increasing the age at which conditions to be classified as disabled are relaxed in the Austrian DI program in 1996. Borghans et al. (2014) analyze the effect of medical reexamination of existing DI recipients based on a more generous criteria as one of the measures introduced by the major DI reform implemented in the Netherlands in 1993. Autor et al. (2016) analyze the effect of a policy change that expanded the medical eligibility criteria in the disability compensation program for veterans in the US in 2001. Mullen and Staubli (2016) analyze the effects of various reforms that changed the generosity of the DI benefits in the Austrian DI program from 1988 to 2004.

[3]The WIA reformed both the disability scheme and the sickness scheme that precedes it. We analyze the total effect of the WIA reform, including the reforms implemented in the two schemes, on work and benefit claiming behavior during the disability period. We can not, and do not, disentangle the impact of each reform on behavior during the disability period.

Hullegie and Koning (2018) analyze the effects of mandating the employer to strengthen its sickness monitoring obligations and pay for the sickness benefit as two measures introduced by the DI reforms implemented in the Netherlands in 2002 and 2004, respectively.

We add to the literature with three main findings from the WIA reform. First, the reform tightened the eligibility criteria much more than earlier reforms did, such as the reform studied by Borghans et al. (2014). Under a much stricter DI regime, sick individuals might be expected to struggle to compensate lost DI benefits, as they might find it harder to increase earnings due to, for example, worse health status. We show, however, that on average sick individuals overcompensate lost DI benefits by increasing earnings and claims from UI. They also increase labor participation although to a lesser extent. Mullen and Staubli (2016) provide supporting evidence that a lower DI replacement rate decreases DI claiming rate but individuals remain able to compensate their lost benefits. Our finding contributes to the growing recognition that even people with severe impairments can work up to some extent (Burkhauser et al., 2014).

Second, we find that the effects of the reform vary substantially with the individual's age, work status, type of employment contract and occupation when falling sick. Hullegie and Koning (2018) find that reforms implemented by the Dutch government in SI have substantially improved work resumption among employees. We extend this finding by differentiating between employees and unemployed, and among employees between those with permanent and temporary contracts, not only for the reforms in SI but also DI. We show that while individuals on a permanent contract compensate €56 of lost DI benefits by increasing earnings by €125 and income from UI by €9, those on a temporary contract compensate €40 of lost DI benefits by increasing earnings by €61 and income from UI by €33. On the other hand, unemployed individuals compensate €51 of lost DI benefits by increasing UI claims by €70, but at the same time they lose a substantial amount of €119 in earnings. The fact that the impact of the reform is very different across labor market groups is of high relevance to policy makers who have to consider the risk of a widening gap between labor market segments as a consequence of welfare reforms.

Furthermore, we find substantial age heterogeneity. Earlier studies provide evidence that DI reforms can both induce work resumption and participation in alternative benefit programs. The reform effects, however, appear to be age dependent. Borghans et al. (2014) find that due to the 1993 DI reform in the Netherlands, individuals around 45 years on average offset €1.00 of lost DI benefits by collecting €0.62 from earnings and €0.30 from other social assistance programs, but this benefit-substitution effect declines over time. Staubli (2011) shows that the DI reform that tightened eligibility criteria for men ages 55-56 years in Austria increased labor participation but also participation in SI and UI. Karlström et al. (2008) study the effect of the DI reform that abolished generous eligibility rules for people 60-64 years old in Sweden. They find that the reform increased participation in SI and UI but not labor participation during the 2-3 years following the reform. These evidences suggest that for older individuals social support substitution is the most likely outcome, whereas for younger individuals work resumption is more likely. While these studies carry out analyses on specific age groups in different countries, we analyze the impact of the reform over a wide range of age groups and investigate the age gradient of the reform that affected sick people of all ages. Our findings over six age groups are in line with the predictions on different age groups from these studies. Following the WIA reform, while individuals over 55 years of age very much struggle to increase earnings and depend on income from UI, younger individuals are much more able to increase earnings and depend less on income from UI.

In addition, we find that responses to the DI reform vary with occupation type. Earlier studies analyzing the effects of DI reforms lack evidence on the role of occupation. We show that all occupation groups increase labor participation in response to losing DI, but they differ

substantially in their ability to compensate for lost DI. While blue-collar workers compensate with earnings and unemployment benefits together, white-collar workers overcompensate with earnings only. Workers who are employed through temporary work agencies fail to compensate with earnings or unemployment benefits.

Third, we find the effects of the reform to be persistent and only slightly diminish over time. This is an important finding from a policy perspective because earlier reforms were less successful in the long run. For example, Koning and Lindeboom (2015) mention that while the DI reforms in the early 1990s at first seemed effective, their impact did not persist because, among other reasons, financial incentives of the reform to reduce DI use among beneficiaries proved ineffective as declines in public DI use were offset by increases in supplementary private DI use. According to Koning and Lindeboom, these reforms were not politically sustainable and were therefore relaxed soon after implementation. We explain the long-term effectiveness of the WIA reform by its interventions early on during the sickness period when employees are still connected to their employers and not thereafter. For workers with permanent contracts, in particular, the reform facilitates work resumption with their own employer, where reintegration prospects are strongest. In other words, in the former DI scheme (WAO), for many DI beneficiaries their work capacities were not fully utilized because the ties with the old employer were prematurely cut.

The remainder of this paper is organized as follows. Section 2 describes the 2006 Dutch DI reform. Section 3 describes the data. Section 4 presents descriptive analysis of the work and benefit claiming behavior of those who are and who are not affected by the reform. Section 5 describes the identification strategy. Section 6 presents the effects of the reform on work and benefit claiming behavior. Section 7 discusses policy implications and concludes.

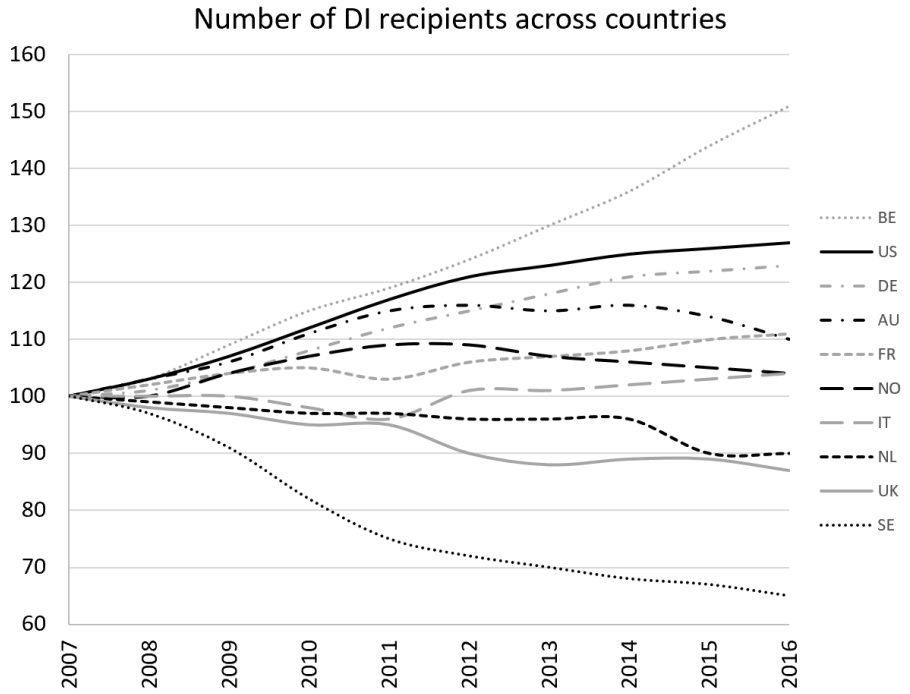

Figure 1: Trends for the number of DI recipients across countries based on the OECD Social Benefit Recipients Database (OECD, 2019). Values normalized (100 in 2007).

2 The 2006 Dutch disability insurance reform

The Disability Insurance Act (WAO) came into effect in 1967 to insure against loss of earnings due to long-term disability. The act was amended several times but since the main amendments in 1993 it preserved its main features until it was replaced by the Work and Income Act (WIA) in 2006. The WAO consists of two schemes. The individual who earns wage or receives UI is first admitted to the sickness scheme if he is unable to perform his work because of illness or injury irrespective of its cause. The duration of the scheme is one year. When the sickness scheme expires, the individual is admitted to the disability scheme if his disability grade is at least 15 percent.[4] The individual is first entitled to the "Wage-loss benefit". When the Wage-loss benefit expires, in the continuation period of the scheme, the individual is entitled to the "Follow-up benefit".[5]

Due to easy access to the WAO, the annual inflow rate into the disability scheme increased to reach 1.5 percent of the insured working population in 2001. Subsequently the disability scheme was reformed. Before the WAO was abolished entirely, however, a transitional scheme was introduced on 1 October 2004 for people who have fallen sick during the period from 1 October 2003 until 1 January 2004. In the transitional scheme the features of the sickness and disability schemes of the WAO have been preserved except that the criteria to enter the disability scheme have been made stricter. In particular, the transitional scheme has adapted a broader definition of what work can still be done by the applicant. This makes it easier to find potential jobs the individual can still perform. As a result, the wage loss due to disability can be estimated to be smaller making it difficult to reach the minimum disability grade to become eligible for DI, or to reach a higher disability grade that leads to a higher Wage-loss benefit.

The WIA came into effect on 1 January 2006 for people who have fallen sick from 1 January 2004 onwards. It introduced major changes in both the sickness and disability schemes to facilitate work resumption and succeeded to limit the yearly inflow rate into the disability scheme to 0.5 percent of the insured working population during the first six years since its introduction (Koning and Lindeboom, 2015). For the sickness scheme, the duration of the scheme is extended from one to two years. The strong incentive for the employer to facilitate work resumption is that it is obliged to compensate the employee for wage loss during the two-year period of the scheme.[6]

For the disability scheme, the WIA inherited the stricter eligibility criteria of the transitional WAO that uses the broader definition of what work can still be done by the applicant. In addition, it introduced a number of major changes. First, the minimum grade of disability required to enter the scheme was increased from 15 to 35 percent. Therefore, workers with limited disability are expected to resume working with adaptations, or to apply for UI.

Second, the scheme introduced a distinction between full and partial disability, and accordingly two specialized disability schemes. If the wage loss is more than 80 percent and there is no potential for any degree of recovery, the worker is admitted to the Full Invalidity Benefit

[4]Disability grade is determined by dividing the estimated wage loss due to disability by the former wage, where estimated wage loss is given by the difference between the former wage and the potential wage that the sick individual can still earn. An ergonomist determines the potential wage by taking the average of the highest wages the sick individual could still earn in three suitable occupations.

[5]The employer is responsible to pay 70 percent of the former wage in the sickness scheme. Most employers pay the full amount. The Wage-loss benefit replaces 70 percent of the former wage multiplied by the disability grade. The duration of the benefit depends on the age of the individual and is limited to a maximum of 6 years. The Follow-up benefit pays the minimum wage and an additional amount that depends on the former wage and the age at which the individual has become entitled to the benefit. The benefit is paid as long as the individual is disabled but expires when he reaches the state pension age.

[6]The compensation must amount to 70 percent of the former wage. Most employers pay the full amount during the first year of sickness, and many pay more than 70 percent of the former wage during the second year.

Regulation (IVA), and is entitled to a benefit that replaces 75 percent of the former wage. Admission to the scheme is limited to a selective group of impairments that are expected to be permanent so that moral hazard is unlikely.

If the wage loss is more than 35 percent and less than 80 percent, or if the wage loss is more than 80 percent but there is still potential for recovery, the worker is insured under the Return to Work Regulation (WGA). The eligible worker is first entitled to the "Wage-related benefit". It replaces 70 percent of the former wage multiplied by the disability grade if the individual utilizes his remaining work capacity to its full potential. The benefit has an UI component that compensates the individual if he is not able to utilize his remaining work capacity.[7] When the Wage-related benefit expires, in the continuation period the individual is entitled to one of two types of benefits. If he utilizes at least 50 percent of his remaining earning capacity, he is entitled to the "Wage-supplement benefit". Otherwise he is entitled to the Follow-up benefit.[8] At a given disability grade, the difference between the Wage-related benefit and the Follow-up benefit is as large as 70 percent of the difference between the former wage and minimum wage, giving the partially disabled workers with higher former wages a stronger incentive to utilize at least 50 percent of their remaining work capacity when the Wage-related benefit expires. Since the UI component of the Wage-related benefit is also exhausted when the Wage-related benefit expires, the individual faces an additional incentive to exploit remaining work capacity in the continuation period of the DI scheme. These mean that, compared to the WAO, the WIA provides stronger incentives to exploit remaining work capacity during the disability period.

Finally, the WIA amended the experience rating in DI. According to experience rating, firms with high disability costs are punished with a higher premium. In the WAO, experience rating applied to employer contributions to DI for a period of 5 years for all disabled workers. In the WIA, the experience rating period is extended to 10 years and applied to employer contributions for partially disabled workers so that the employer is financially incentivized for a longer period to reintegrate beneficiaries with remaining work capacities.

In the WAO or WIA, if the individual has no employer during his participation in the sickness scheme, he is eligible for the "Sickness benefit" (ZW) that replaces 70 percent of the former wage. During participation in the disability scheme, the individual is eligible for the UI.[9] The amount of the UI is a certain fraction of the remaining earning capacity. In the WAO, the individual is required to file an application to claim the UI. In the WIA, however, the UI is integrated into the DI, and therefore no application is required. In fact, the duration of the Wage-related benefit is determined by the duration of the UI.

3 Data

We use unique administrative data from the Employee Insurance Agency (UWV) on three cohorts of sick people who face different criteria to enter the disability scheme and different incentives to resume working if participating in the disability scheme. In particular, the data contains information on all individuals who fell sick in the third quarter of 2003, fourth quar-

[7] The duration of the benefit depends on the employment history, and is limited to a maximum of 38 months.

[8] The Wage-supplement benefit replaces 70 percent of the former wage multiplied by the disability grade. The Follow-up benefit replaces 70 percent of the minimum wage multiplied by the disability grade. Both benefits are paid as long as the individual is disabled but expire when he reaches the state pension age.

[9] If the benefit received from a benefit scheme (sickness, disability, or unemployment scheme), or the wage earned during the second year of sickness (in the WIA), is lower than the applicable social minimum, it is supplemented up to the social minimum according to the Supplementary Benefits Act (Toeslagenwet). The total of the benefit and the social minimum supplement cannot exceed the former wage. If the individual is living with a partner, the supplement is granted if the total income of the individual and the partner is below the social minimum. If the individual is living alone, the amount of the supplement depends on whether the individual has children.

ter of 2003, and the first quarter of 2004, and therefore became eligible to participate in the WAO, transitional WAO, and the WIA schemes, respectively. For these people we observe the beginning and ending dates of sickness, gender, date of birth and industry of main activity. These people either earn wage or receive UI at the time they fall sick since people of other labor market groups are not eligible to enter the sickness scheme. For people in employment, we observe whether they hold a permanent contract, temporary contract or a contract through a temporary work agency.

We merge the administrative data on sickness with administrative data on labor participation, wage, and benefits, all available on a monthly basis from Statistics Netherlands. The benefits are from various benefit schemes which include DI, UI, general assistance for low-wage earners, and other benefits from a large number of smaller benefit programs. The data from Statistics Netherlands extend from January 1999 to December 2015, and allow to study the differences in benefit claiming and work behavior of the three cohorts of sick individuals over a long period of time.

The initial sample of sick people consist of 51,319,668 observations for 251,567 individuals. We impose three restrictions on the initial sample. First, we restrict the sample with respect to the number of days spent in sickness. Employers are mandated to report to UWV sickness cases if they last longer than 90 days. Temporary work agencies are the main suppliers of short-term sickness cases. They complied with mandatory reporting to a large extent only from January 2004 onwards, after the WIA has come into effect. Therefore, in the data, short-term sickness cases are under-reported for participants of the (transitional) WAO who fell sick before January 2004. The three cohorts of sick individuals, however, share similar distributions of sickness duration when sickness is restricted to last at least 180 days. Therefore, we drop the sickness cases that last shorter than 180 days.[10] This restriction leads to a sample of 19,690,488 observations for 96,522 individuals. Second, we drop the individuals who start receiving DI before they become eligible to apply for DI. The majority of this group are existing DI recipients due to an earlier sickness application and assessing the impact of the reform is much harder among the members of this group as they participate in multiple schemes or participate in the same scheme multiple times. This restriction leads to a sample of 16,028,076 observations for 78,569 individuals. Third, we drop individuals if they are participants of the disability schemes for the self-employed (WAZ) or young people (WAJONG) since the institutional rules and incentives for work resumption are very different for them. This restriction leads to a sample of 15,864,876 observations for 77,769 individuals which constitute the study sample.

As described above, sick individuals become eligible to participate in one of three disability schemes depending on the date they fall sick. This allows to construct control and treatment groups and compare their responses to the disability reform in a quasi-experimental research design. In particular, we categorize the sick individuals into three groups based on the date they fall sick: the WAO group, the transitional WAO group, and the WIA group, which consist of individuals who fell sick in the third quarter of 2003, fourth quarter of 2003, and first quarter of 2004, respectively. We consider the WAO group as the "control group", and the transitional WAO or the WIA group as the "treatment group".

Based on the available data on labor and benefits, two sets of five outcome variables are defined and used to compare the behavioral responses of control and treatment groups to the

[10] Figure A.1 presents distributions of sickness duration for individuals who fall sick in select months and participate in the WAO, transitional WAO, and the WIA. The left panel presents distributions for the three groups when the number of days spent in sickness ranges from 0 to 360 where 360 is the the maximum number of days individuals can spend in the sickness scheme if they participate in the WAO or the transitional WAO. The right panel presents distributions when the number of days spent in sickness ranges from 180 to 360. In Appendix A we show that regression results from samples based on alternative numbers of days spent in sickness are largely in line with the baseline regression results.

disability reform. The first set considers labor participation and benefit receipt, and the second set considers income from labor and benefit programs. As described in Section 2, UI is integrated into the DI in the WIA while this is not the case in the WAO. With respect to DI, an outcome variable of benefit receipt is defined to indicate receipt of DI and possibly also UI at the same time. This means that while the institutional setting requires that the individual who receives the DI also receives the UI at the same time if he is insured under the WIA, in our definition of benefit receipt, we allow that the individual who receives the DI can also receive the UI if he is insured under the WAO. An outcome variable of income is defined to indicate income from DI and possibly also from UI. Again, while the DI includes the UI if the individual is insured under the WIA, in our definition of income, we add the UI to the DI if the individual is insured under the WAO and receives the UI in addition to the DI. With respect to UI, an outcome variable of benefit receipt is defined to indicate receipt of UI but not the DI at the same time. An outcome variable of income is defined to indicate UI exclusive of DI. With respect to general assistance and other benefits, indicators of receiving benefits, and income from these schemes are defined as outcome variables. With respect to employment, labor participation and wage income are defined as outcome variables.

4 Descriptive analysis

Time trends in outcome variables

Figures 2a and 2b show the time profiles of labor participation and benefit receipt, and the time profiles of income from work and benefit programs, for control and treatment groups over a period of 16 years from January 1999 to July 2015. A time profile of a given group is generated as follows. First, within a group and in a given calendar month, the mean of the outcome variable (dummy that indicates labor participation or receipt of a benefit, or income from work or benefits) is calculated. The set of means calculated for each month of the 16-year period are then used to draw a time profile. In the subfigures vertical lines are added at the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO, and WIA schemes.

The top left subfigure in Figure 2a shows the time profile of probability of receiving DI (possibly also UI at the same time). A large fraction of about 20 percent of the sick individuals insured under the WAO and transitional WAO claim DI at the onset of the disability period. This fraction shows no notable change until the end of the study period. The time trends of the WIA and WAO groups are similar, except in two respects. First, the inflow into the disability scheme for the WIA group is not as immediate as it is for the WAO groups when individuals become eligible to apply for DI. It might be that stricter rules of the WIA make it difficult to claim DI at the first attempt. Second, the time trends for the WAO groups show a decrease during the first year after they show a peak at the onset of the disability period. It might be that the WAO groups have better chances of recovery shortly after entering the disability scheme. The WAO groups are relatively healthier than the WIA group since in the WIA the minimum disability grade to enter the disability scheme is higher but also disability assessment is stricter.

The probability of working shows a strong time trend that is common to all groups. It increases until the date individuals fall sick. This pattern does not reflect behavioral responses, but it reflects the fact that individuals can enter the sickness scheme, and get reported as sick in the administrative data, only if they are working or receiving UI at the time they fall sick (Section 3). Before this time, these individuals can have another status outside the labor force. The probability of working decreases dramatically during the first year of sickness, remains

fairly stable for about three years, but decreases further throughout the remaining years of the study period.

UI use appears to be strongly related to the use of sickness and disability benefits. For the WAO groups, during the sickness period, UI use decreases sharply because unemployed people who fall sick change their UI for sickness benefit. UI use rebounds thereafter because many of these people recover during the sickness period. It peaks as individuals become eligible to apply for DI because when the sickness period ends, those who apply but get rejected to enter the disability scheme turn to UI. UI use decreases during the disability period. This is because UI is temporary and ends after a maximum of 38 months. For the WIA group a similar time trend is observed except that UI use increases further during the second year of sickness before it peaks in the beginning of the disability period since members of this group have more time for recovery during the longer sickness period and change their sickness benefit for UI.

The probability of receiving general assistance increases steadily from the time individuals fall sick. A similar but less pronounced increase is observed for the probability of receiving other benefits from various small benefit programs from the time individuals fall sick. These time patterns appear to be related to that of working. As individuals work less due to sickness, their earnings decrease enough to become entitled to these benefits.

In the right panels of Figures 2a and 2b, we show the time profiles of the means of income from work and benefit programs. In a mean calculation, we allow for zero income to analyze how the reform affected the income received due to changes in the amount of wages earned and benefits received at the intensive margin, but also due to participation in the labor market and use of benefit programs at the extensive margin. The time trends of the means of different types of income resemble those of participation in the labor market and use of benefit programs shown in the left panels of the figures.

Differences between the control and treatment groups

Our aim is to compare the behavioral responses of control and treatment groups before these groups fall sick and after they become eligible to apply for DI. Here we conduct exploratory graphical analysis of the differences between the control and treatment groups in the outcome variables during these two periods to provide a first evidence on the impact of the reform. Based on Figures 2a and 2b, comparing the time profiles of a given outcome variable across the control and treatment groups is, however, difficult because they often exhibit strong trends, while they are close to each other at the same time, as in the subfigure for wage income in Figure 2a.

To make an explicit comparison of the behavioral responses across the control and treatment groups, in Figures 3a and 3b we draw the time profiles of the outcome variables where we net out the effect of calendar time. This is done as follows. We obtain residuals from a regression where labor participation or benefit receipt, or income from labor or a benefit, is the outcome, and a full set of calendar month dummies for the entire study period and time-invariant individual fixed effects are controls. Within each group (control or treatment) and at given calendar months, we compute the average of the residuals. We use the computed set of mean residuals across all calendar months of the study period to draw the time profile of the residualized outcome variable for each group. If the outcome variable is, for example, the residualized participation in labor or benefit receipt, a given point on the time profile of a given group can be interpreted as the propensity of that group to participate where calendar time plays no role. In the figures we do not present any pattern during the sickness period as our interest lies in comparing behavioral responses across the control and treatment groups before individuals fall sick and after they become eligible to apply for DI.

The main observation is that, in all subfigures of Figures 3a and 3b, the differences between

the WIA and WAO groups are statistically insignificant before individuals fall sick, while they are substantial and significant when individuals are eligible for DI. The stricter WIA regime appears to lead to the following differences between the WIA and WAO groups. The WIA group is much less likely to receive DI, especially during the first years after all groups become eligible to apply for DI. The gap between the WIA and the WAO groups remains at about 5 percentage points until the end of the study period, providing descriptive evidence of the magnitude of the overall impact of the reform on DI use. As the WIA group is less likely to make use of DI, the group appears more likely to work than the WAO groups during the entire period individuals are eligible for DI. The size of the gap between the WIA and WAO groups is not particularly large, however. It might be that labor supply responses to the stricter WIA regime often occur at the intensive margin rather than the extensive margin. Another possibility is that the WIA group seeks participation in alternative benefit programs. In fact, this group is more likely to receive UI.[11] On the other hand, this group is less likely to claim general assistance or other benefits. A potential reason is that the higher wages earned by the WIA group makes this group less likely to fall below the income threshold to become eligible for means-tested or subsistence level benefits. Through the end of the study period, however, the WIA group becomes less likely to receive UI possibly because the duration of the UI expires. Accordingly, the WIA group becomes as likely as the WAO groups to claim general assistance and other benefits.

The differences between the transitional WAO and WAO groups are often insignificant before individuals fall sick or after they become eligible to apply for DI. A notable exception is that, for a period of about two years since individuals become eligible to apply for DI, the probability of receiving DI is smaller for the transitional WAO group than for the WAO group. This is expected because the stricter eligibility criteria of the transitional WAO scheme makes it more difficult to reach the minimum disability grade to become eligible for DI (Section 2). However, after this period, the gap between the two groups narrows and becomes statistically insignificant. This can be explained by the reexaminations of the WAO participants younger than 45 years of age that have taken place from 2004 until 2008 based on the eligibility criteria of the transitional WAO scheme (Garcia-Mandicó et al., 2020).

The time profiles of residualized income outcomes, presented in the right panels of Figures 3a and 3b, show that the WIA group receives smaller amounts of DI and general assistance, and earns substantially more wage during the disability period. The differences for other income outcomes are much smaller or insignificant. The differences between the groups are small and insignificant during the period before individuals fall sick in all outcomes of income.

[11] The spikes for the transitional WAO and WIA groups in the subfigures for working, earnings, UI receipt, and income from UI are due to that the sickness data is for individuals who earn wages or receive UI (Section 2), and these groups did not fall sick yet at the time of the spikes while the WAO group did.

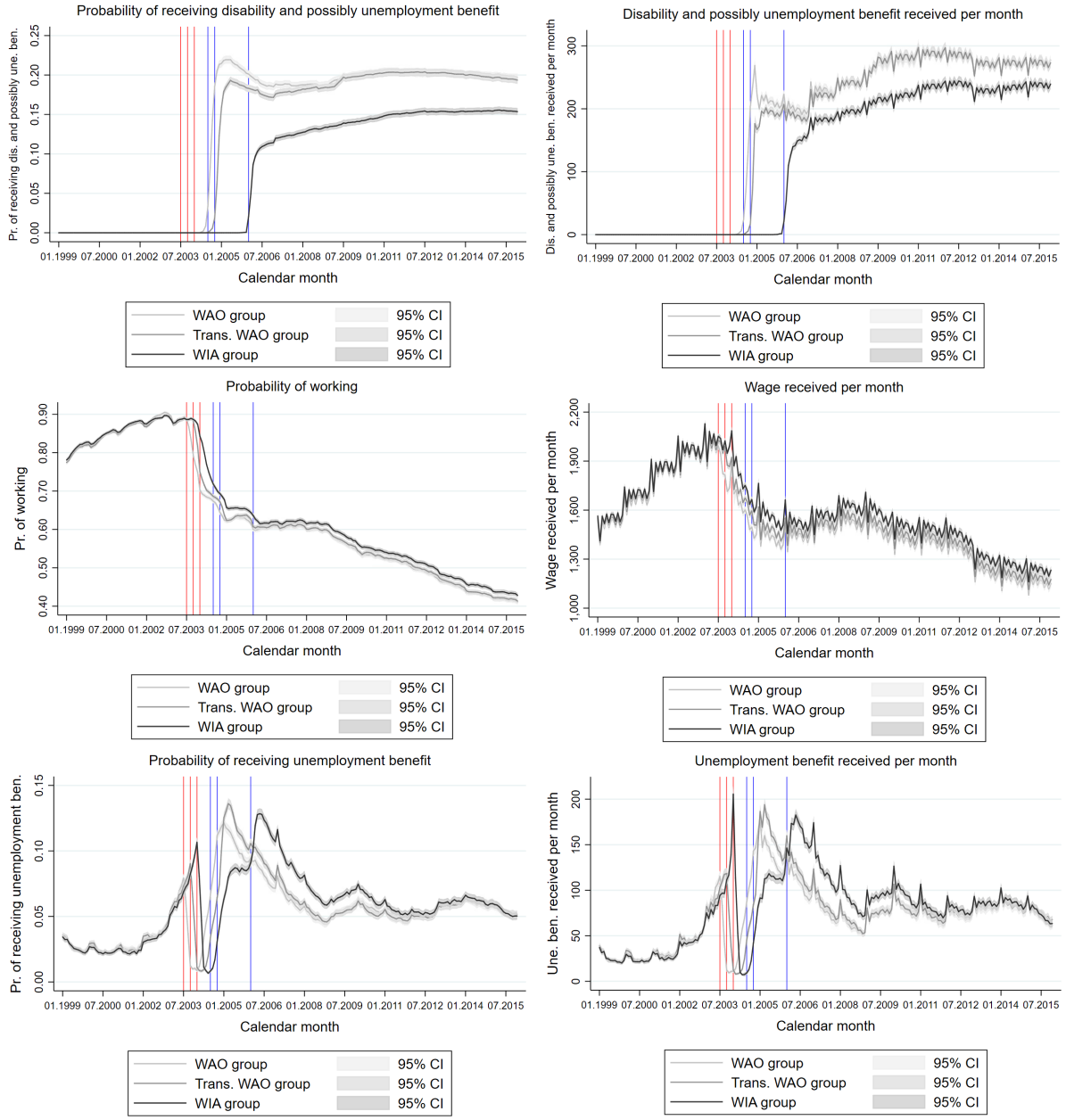

Figure 2a: Labor participation, benefit (DI and UI) receipt, labor income and benefit income for control and treatment groups over calendar months. In a given subfigure, a point on a given time profile represents the mean of the outcome variable (dummy that indicates labor participation or benefit receipt, or income from work or benefits) within a group in a given calendar month. Around the mean is a 95 percent confidence interval. DI might be supplemented with UI. UI is defined so that receiving DI at the same time is not allowed. Each subfigure is based on the study sample of 15,864,876 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

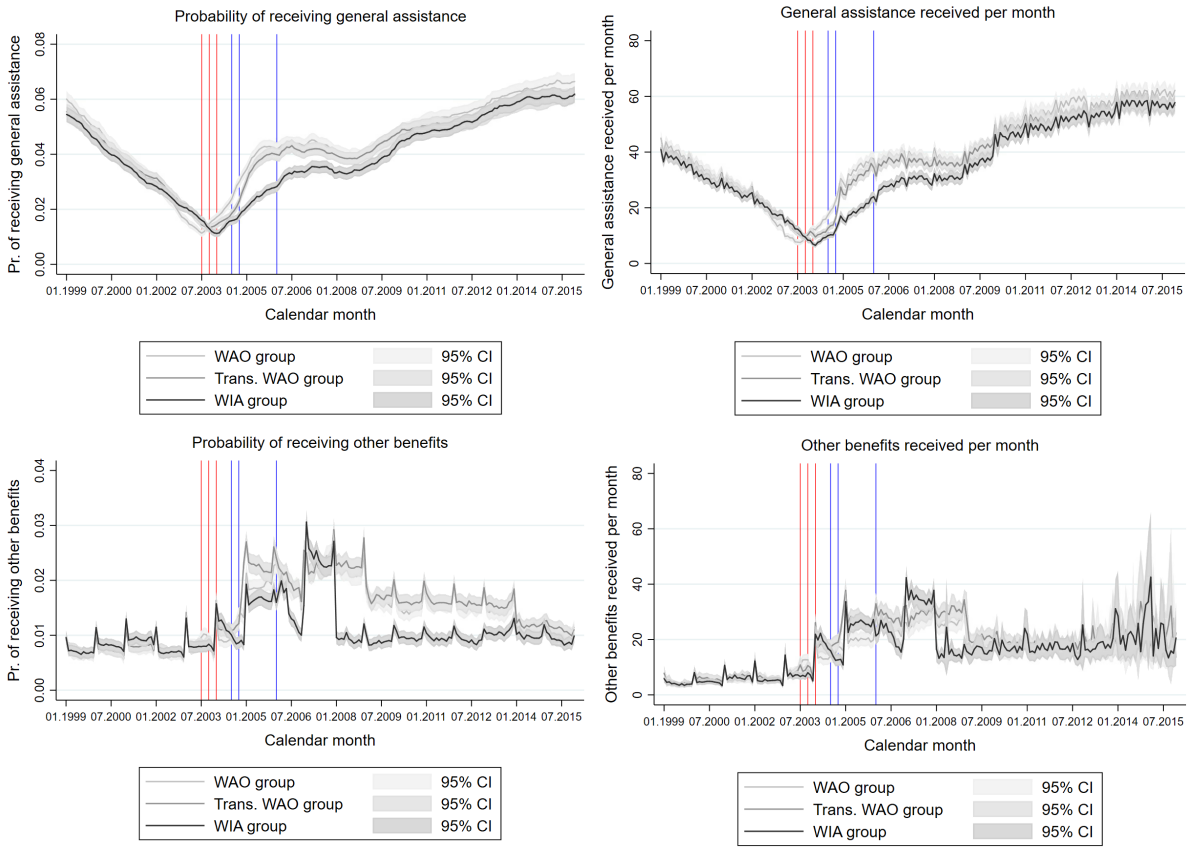

Figure 2b: Benefit (general assistance and other benefits) receipt and income for control and treatment groups over calendar months. In a given subfigure, a point on a given time profile represents the mean of the outcome variable (dummy that indicates receipt of a benefit, or income from benefits) within a group in a given calendar month. Around the mean is a 95 percent confidence interval. Each subfigure is based on the study sample of 15,864,876 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

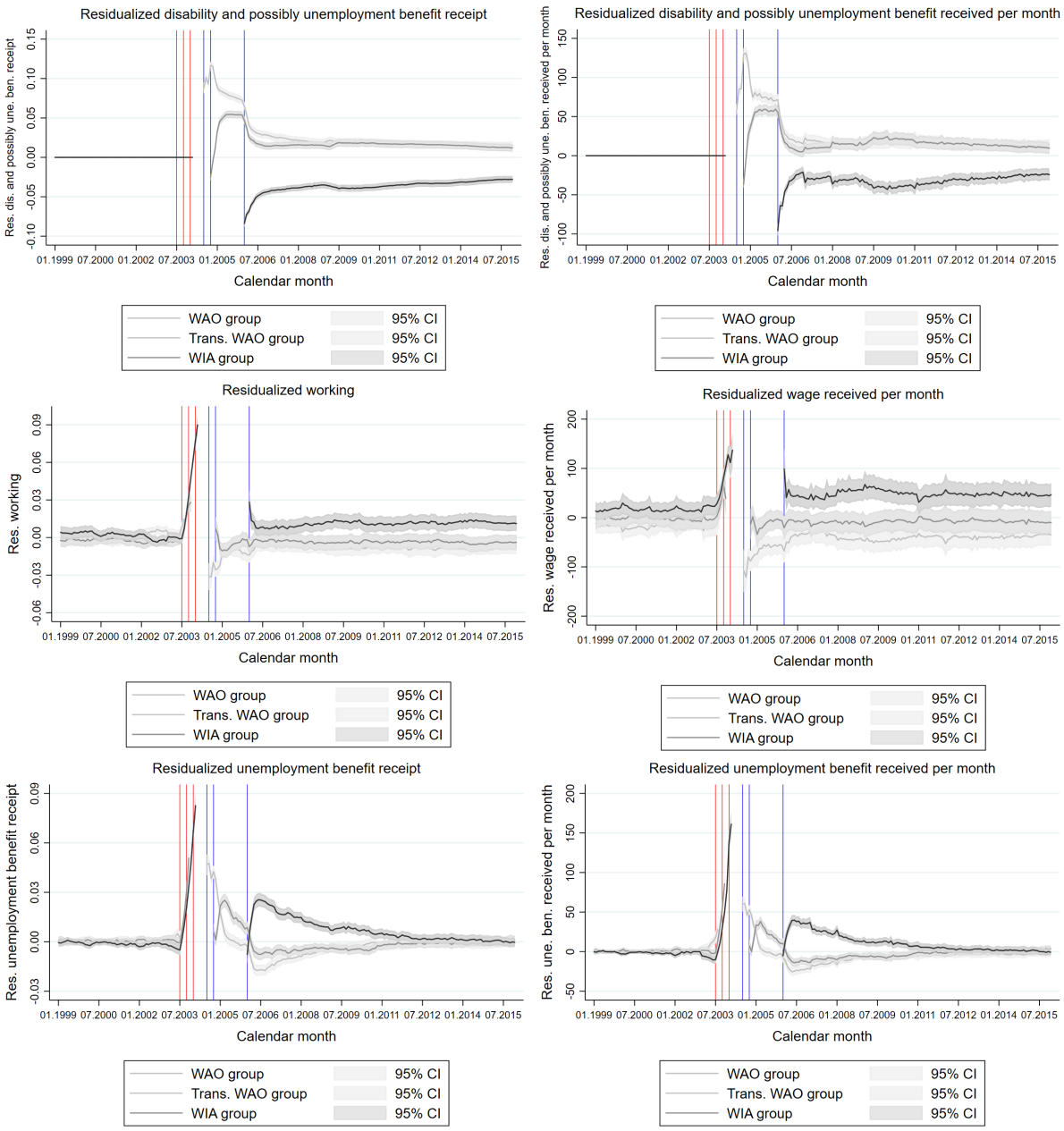

Figure 3a: Residualized labor participation, benefit (DI and UI) receipt, labor income and benefit income for control and treatment groups over calendar months. In a given subfigure, a point on a given time profile represents the mean of residuals within a group in a given calendar month. Around the mean is a 95 percent confidence interval. Residuals are from a regression where labor participation, benefit receipt, labor income, or benefit income is the outcome, and a full set of calendar month dummies for the entire study period and time-invariant individual fixed effects are controls. DI might be supplemented with UI. UI is defined so that receiving DI at the same time is not allowed. Time profiles are not considered during the sickness period. Each subfigure is based on 4,509,329 and 10,117,027 observations during the periods that precede and succeed the sickness period, respectively, for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

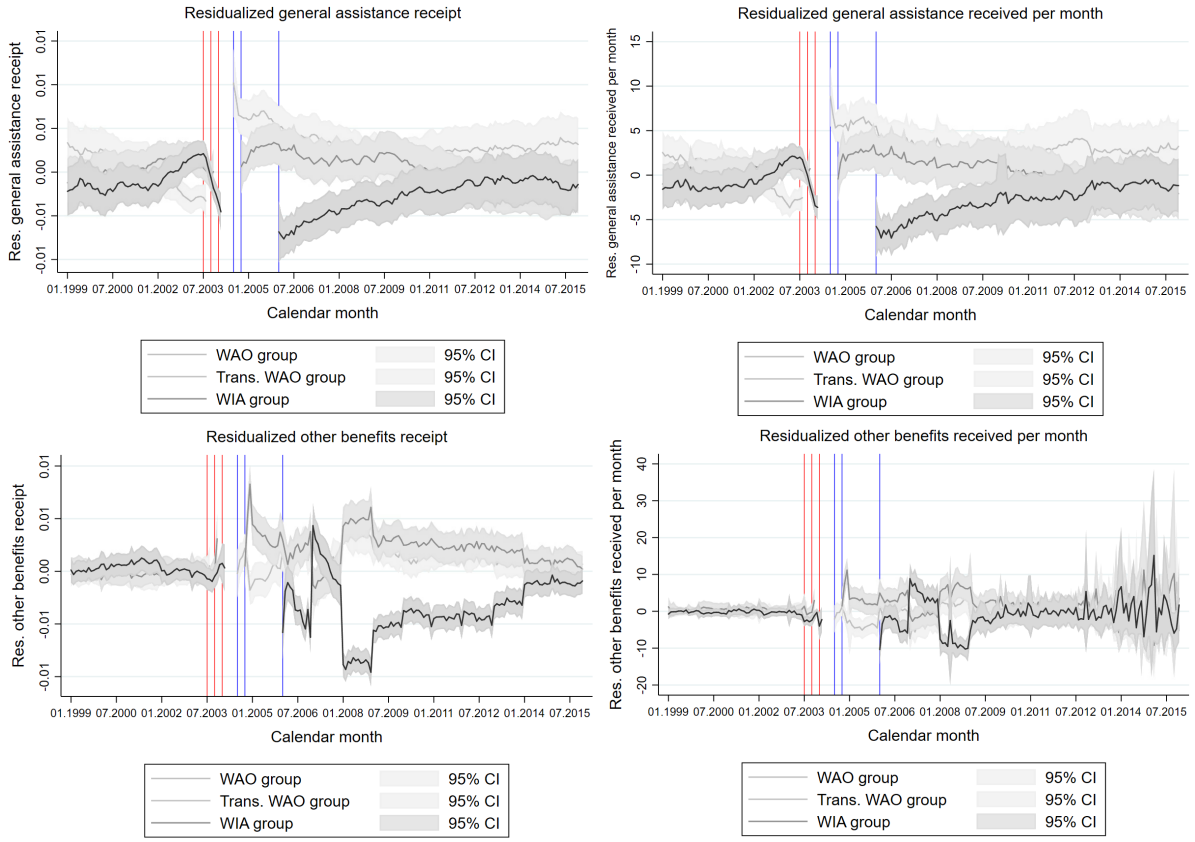

Figure 3b: Residualized benefit (general assistance and other benefits) receipt and income for control and treatment groups over calendar months. In a given subfigure, a point on a given time profile represents the mean of residuals within a group in a given calendar month. Around the mean is a 95 percent confidence interval. Residuals are from a regression where benefit receipt or income is the outcome, and a full set of calendar month dummies for the entire study period and time-invariant individual fixed effects are controls. Time profiles are not considered during the sickness period. Each subfigure is based on 4,509,329 and 10,117,027 observations during the periods that precede and succeed the sickness period, respectively, for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Vertical lines indicate the first instance individuals could become entitled to the sickness (red) and disability (blue) benefits in the WAO, transitional WAO and WIA schemes.

5 Identification strategy

We use a difference-in-differences (DiD) approach to identify the causal effects of the transitional WAO and the WIA reforms on labor participation and benefit receipt, and on income from labor and benefit programs. The first difference is across groups. Those who fall sick in the last quarter of 2003 or the first quarter of 2004 are subject to different eligibility criteria to access DI, and they face different incentives to work or claim benefits, compared to individuals who fall sick in the third quarter of 2003 and are subject to a less restrictive benefit regime. The second difference is over (event) time. After individuals become eligible to apply for DI, the incentives to work or claim benefits change compared to the period before individuals fall sick.

We implement the DiD comparison using the following regression:

$$y_{it} = \gamma_1 \left(Treat_i^{Trans.\ WAO} \times Post_t\right) + \gamma_2 \left(Treat_i^{WIA} \times Post_t\right) + \delta\, Post_t + \lambda_s + \alpha_i + \varepsilon_{it}. \quad (1)$$

i indexes individuals. t indexes the months of event time. Event time indicates the period of time before individuals fall sick, and subsequently the period of time when individuals are eligible for DI. t indexes the months of the former period with values from –55 to 0, and those of the latter period with values from 1 to 120. y_{it} is the outcome variable of interest. λ_s is a monthly calendar time effect. s indexes the calendar months of the period before individuals fall sick, and subsequently the calendar months of the period when individuals are eligible for DI.[12] α_i is an individual-specific, time-invariant intercept term. ε_{it} represents the individual-specific, time-varying shocks that are not observed. These error terms are assumed to be uncorrelated among each other and with all the explanatory variables.

$Treat_i^{Trans.\,WAO}$ and $Treat_i^{WIA}$ are dummy variables that indicate the treatment groups, in particular the transitional WAO and the WIA groups, respectively.[13] The periods of time before individuals fall sick and after they become eligible to apply for DI are the pre-treatment and post-treatment periods, respectively. $Post_t$ is a dummy variable that indicates the post-treatment period. We interact $Treat_i^{Trans.\,WAO}$ and $Treat_i^{WIA}$ with $Post_t$ to capture the mean difference in the outcome variable between the treatment and control groups during the post-treatment period compared to the mean difference between the two groups during the pre-treatment period. In this comparison, the latter difference aims to account for differences between the groups due to factors other than the policy reform. γ_1 and γ_2 are the coefficients of main interest and reflect the effects of the transitional WAO and WIA reforms, respectively. The assumptions needed to obtain unbiased estimates of these coefficients are discussed below. Standard errors of the coefficient estimates are adjusted for clustering at the individual level.

It is important to note that we interpret that the differences in behavior between the treatment and control groups during the disability period are not only due to the stricter eligibility criteria for disability benefits (disability reform) but also due to the extension of the sickness period from one to two years (sickness reform). For example, the extra year spent in the sickness scheme provides the individuals in the treatment group with more time for recovery and might reduce their propensity to participate in the disability scheme. We cannot and do not aim to separately identify the effects of the different incentives of the sickness and disability reforms on work and benefit claiming behavior during the disability period. Instead, we analyze the total effect of the WIA reform, comprising of the sickness and disability reforms, on behavior during the disability period, relative to the period before individuals fall sick.

[12] For example, for the individual who falls sick in July 2003 and participates in the WAO scheme, where the sickness period lasts one year, s indexes the months from January 1999 to February 2004, and subsequently the months from July 2004 to July 2015. January 1999 is chosen as the base month for comparison.

[13] These group dummies have no time variation and are omitted in the fixed effects regression.

To study the effects of the transitional WAO and WIA reforms in the short and long run, we consider the following regression:

$$y_{it} = \sum_{l=1}^{10} \gamma_{1l} \left(Treat_i^{Trans.\,WAO} \times d_{lt} \right) + \sum_{l=1}^{10} \gamma_{2l} \left(Treat_i^{WIA} \times d_{lt} \right) + \sum_{l=1}^{10} \delta_l d_{lt} + \lambda_s + \alpha_i + \varepsilon_{it}. \tag{2}$$

Compared to Equation (1), instead of the $Post_t$ dummy, which indicates the entire post-treatment period, we consider year dummies. That is, we divide the post-treatment period into ten equal periods of one year each. d_{lt} is a dummy that indicates a one-year period of the post-treatment period. For example, d_{2t} indicates the second year from the time the individual becomes eligible for DI. The pre-treatment period is chosen as the base period for comparison. The interaction terms of treatment and year dummies capture the mean difference in the outcome variable between the treatment and control groups in a given year compared to the mean difference between the two groups in the pre-treatment period.

In the described DiD setup, treatment and control groups are compared over "event time". Therefore, a distinction is made between event time and "calendar time". That is, according to our institutional setting, members of the treatment and control groups fall sick at different points in time, but also the durations of the sickness schemes they participate in are different, which both lead to different calendar dates to become eligible to apply for DI. However, our aim is to compare the behavior of treatment and control groups across the periods before these groups fall sick and after they become eligible to apply for DI. This means that our DiD comparison requires comparing members of the treatment and control groups at different points in time with respect to the calendar months these groups fall sick but also the calendar months they become eligible to apply for DI. Therefore, "event time" is defined to refer to the time periods before individuals fall sick and after they become eligible to apply for DI. "Calendar time" is defined to define calendar time effects that aim at capturing calendar time trends in outcome variables. In Equation (2), the year dummies of event time are correlated with the month dummies of calendar time by construction. This correlation structure could affect the point estimates or standard errors of the main effects of interest given by the coefficients of the interaction terms. Results based on three-year period dummies with much weaker correlations with calendar month dummies show that estimates of the effects of the interaction terms are very similar to those when year dummies are used so that our qualitative results are not affected (see Appendix A).

Are the treatment effects likely to be confounded by unobserved shocks?

In Equation (1), to obtain unbiased estimates of the effects of the transitional WAO and WIA reforms, it must hold that the treatment and post interactions are not correlated with economic shocks captured by the error term of the regression. According to the institutional setting, falling sick in one of three adjacent quarters of a nine-month period determines eligibility for one of three disability schemes. As eligibility for different schemes is determined within a short period of time, economic shocks should affect the three cohorts of sick individuals in similar ways so that the differences in outcomes across these cohorts can still be attributed to the differences in treatments under different benefit regimes. During the nine-month period from July 2003 to March 2004, however, no particular changes in macroeconomic conditions are observed, or no other major social security reforms are introduced.

Are sick individuals likely to select themselves to disability schemes?

As described in Section 2, falling sick during one of three adjacent quarters of the nine-month period from July 2003 to March 2004 determines eligibility for one of the three disability schemes. This means that individuals may react to select themselves into one of the three disability schemes from the moment the transitional WAO and WIA reforms are announced. For example, they may select themselves into the lenient WAO scheme before it gets abolished at the end of September. Or, they may react in anticipation of the larger WIA reform having seen the transitional WAO reform and select themselves into the transitional WAO scheme. In these cases, the causal effect of the reform cannot be identified. However, self-selection into different disability schemes is unlikely. The government presented a general policy program outlining, among other targets, its plan to reform the WAO scheme on 15 September 2003. They announced that the sickness period will be extended from one to two years, and a stricter DI law will be introduced for the individuals who fall sick from 1 January 2004. The transitional WAO reform was announced on 12 March 2004. The details of the WIA reform were announced on 18 August 2004. This means that for people who face adverse health conditions in September 2003, sick reporting in anticipation of the transitional WAO reform is impossible. However, following the first announcement in September 2003, individuals could report sickness during the last quarter of 2003 instead of after the introduction of the WIA reform on 1 January 2004.

Figure 4 presents the number of individuals by the month they fall sick. Sick reporting shows a decreasing trend in the last quarter of 2003 and shows a notable increase in January 2004. This does not suggest that individuals select themselves into the lenient transitional WAO scheme to avoid participation in the much stricter WIA scheme.

Are the control and treatment groups observationally equivalent?

Table 1 presents sample means and balancing tests of background characteristics and outcomes in control and treatment before individuals fall sick and after they become eligible to apply for DI. Panel A of the table presents the sample means and group mean differences for a number of background characteristics. A sample mean is calculated as the average of a given characteristic of all individuals in a given group at the time these individuals fall sick. In all groups, the average age is about 40 and the fraction of men is slightly higher. About 60 percent hold a permanent work contract, less than 15 percent hold a temporary contract or a contract through a temporary work agency, and about 20 percent is unemployed.

Columns 4 and 5 of the table present group mean differences based on linear regression and indicate whether the difference is statistically significant. In particular, a difference represents the estimated coefficient from the regression of the characteristic as the dependent variable, and the treatment indicator (whether member of the transitional WAO or the WIA group) as the explanatory variable. The differences are often small but statistically significant at the 1 percent level. The reported F-statistic tests whether the estimated coefficients of all characteristics are jointly zero in a regression where treatment indicator is the dependent and all background characteristics are regressors. The statistic is significant at the 1 percent level. The significant group mean differences as well as the F-statistic suggest that background characteristics affect which disability scheme the sick individual participates in. However, background characteristics are not likely to directly affect participation in a particular scheme but affect it indirectly through the differences in the dates individuals fall sick. As described above, sick individuals become eligible to participate in one of three disability schemes depending on the date they fall sick during the period from July 2003 to March 2004. The differences in the dates individuals fall sick are likely to drive the differences in observed background characteristics across the control and treatment groups. For example, the two treatment groups less often work on temporary

contracts but are more often unemployed compared to the control group. It might be that fewer temporary jobs are available in the fourth quarter of 2003 and the first quarter of 2004 when individuals in the two treatment groups fall sick due to a seasonal change in demand for labor. In our DiD regression based on Equation (1), we account for the differences in age, gender and labor market status at the time of falling sick which are invariant over time.

Panel B of Table 1 presents sample means and group mean differences for labor participation and benefit receipt and income from labor and benefit programs. During the pre-treatment period, a number of the group mean differences are significant, especially for the WIA group. These differences are not behavioral. By definition, members of the WIA group fall sick at least three months later than the members of the WAO group. This means that in the calculation of the mean of an outcome variable for the WIA group, additional observations from the pre-treatment period are used leading to a group mean that is higher or lower than that for the WAO group. For example, in Figure 2a, the subfigure for wage income shows that the WIA group earns wage for additional months before reporting sick when the WAO group has already reported sick, which leads to a higher mean wage for the WIA group. In our DiD regression based on Equation (1), monthly calendar time effects account for group specific calendar time trends during the pre-treatment period.

During the post-treatment period, the mean differences are often statistically significant suggesting that the reforms induce behavioral changes. The differences are almost always larger for the WIA group than for the transitional WAO group reflecting the fact that the WIA regime is stricter than the transitional WAO regime. These differences in group means are in line with the differences in the time profiles of group means across the months of the post-treatment period observed in Figure 3a and Figure 3b.

Are the pre-reform time trends common to control and treatment groups?

The main assumption of our identification strategy is that control and treatment groups share common time trends in labor and benefit outcomes before and after these groups fall sick and face the reform incentives. The assumption is testable during the pre-reform period. Figures 2a and Figure 2b show that control and treatment groups share very similar time trends until individuals fall sick, providing visual evidence for this main identifying assumption. Here we use regression analysis to test this assumption. In particular, we use pre-reform data from January 1999 to June 2003 to estimate a regression where labor participation, benefit receipt, labor income, or benefit income is the outcome, and calendar month dummies, interactions of treatment and calendar month dummies, and time-invariant individual fixed effects are controls. January 1999 is chosen as the base month for comparison. Coefficient estimates of interaction terms that are not statistically different from zero provide evidence in favor of the common trends assumption.

Figures 5a and 5b plot the coefficient estimates of the treatment and calendar month dummy interactions from regressions for different outcomes. For both treatment groups, the coefficient estimates are not statistically different from zero throughout the pre-reform period in the subfigure of other benefits. The coefficient estimates deviate from zero to different extents at the end of the pre-reform period in the subfigures of other outcomes. Based on the F-tests presented in the notes below the figures, for these outcomes, we reject the hypothesis that all the interaction terms are jointly equal to zero at the 5 percent level in seven of sixteen regressions. Besides this regression based test on the equality of the time trends of the control and treatment groups, however, the visual evidence based on the time trends in Figures 2a and 2b suggest that control and treatment groups share very similar downward or upward time trends. Furthermore, additional sensitivity analysis shows that the statistically significant differences in the time trends

of treatment and control groups at the end of the pre-reform period have no notable impact on the baseline estimates of the reform effects (see Appendix A).

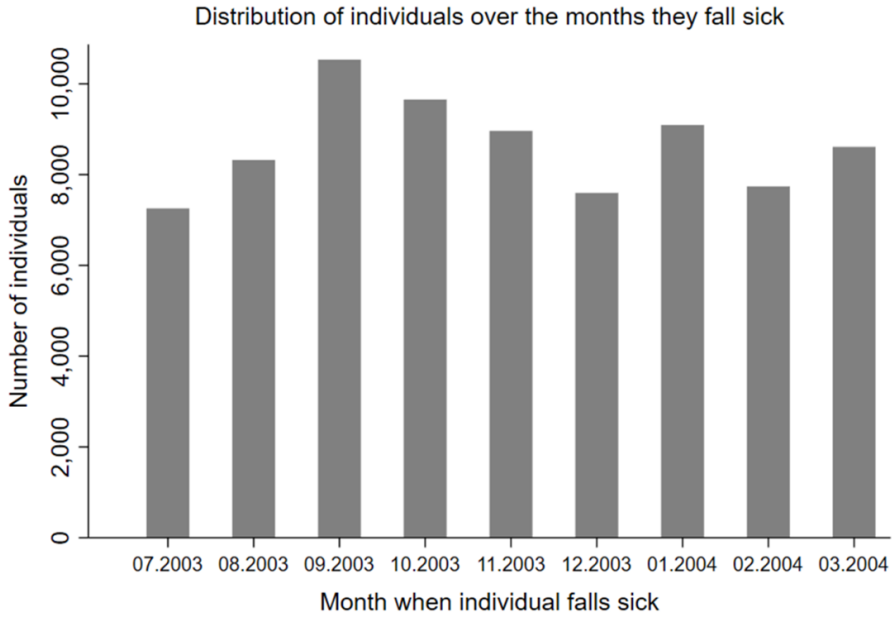

Figure 4: Distribution of individuals over the months they fall sick.

Table 1: Sample means and balancing tests of background characteristics and outcomes in control and treatment before and after the sickness period

	Before					After				
	WAO group	Trans. WAO group	WIA group	Dif. trans. WAO and WAO	Dif. WIA and WAO	WAO group	Trans. WAO group	WIA group	Dif. trans. WAO and WAO	Dif. WIA and WAO
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Background characteristics										
Age	39.690	40.137	40.628	0.447***	0.937***					
Female	0.441	0.437	0.440	−0.004	−0.001					
Permanent contract	0.581	0.574	0.593	−0.007	0.011***					
Temporary contract	0.138	0.130	0.119	−0.008***	−0.020***					
Temporary contract via agency	0.048	0.049	0.034	0.001	−0.014***					
Unemployed	0.188	0.204	0.213	0.015***	0.025***					
Other	0.044	0.043	0.041	−0.001	−0.003					
F-statistic				8.210***	32.75***					
B. Labor market outcomes										
DI (and possibly UI) receipt						0.198	0.191	0.141	−0.007**	−0.057***
Labor participation	0.858	0.856	0.860	−0.002	0.003	0.534	0.536	0.539	0.002	0.004
UI (no DI) receipt	0.033	0.036	0.037	0.002**	0.003***	0.064	0.065	0.069	0.002*	0.005***
General assistance receipt	0.036	0.035	0.032	−0.001	−0.003***	0.050	0.048	0.047	−0.002*	−0.004**
Receipt of other benefits	0.008	0.008	0.008	0.000	0.001	0.017	0.018	0.012	0.001*	−0.005***
DI (and possibly UI) rec. per month						253.895	250.790	212.504	−3.105	−41.391***
Wages earned per month	1768.763	1792.750	1820.622	23.987**	51.859***	1376.214	1407.520	1453.732	31.306**	77.518***
UI (excl. of DI) received per month	40.620	43.528	45.717	2.908**	5.097***	86.192	90.183	96.359	3.991***	10.167***
General assistance received per month	27.657	26.730	25.061	−0.927	−2.596**	47.287	44.936	44.009	−2.351	−3.277**
Other benefits received per month	5.650	6.486	5.922	0.836	0.273	20.346	23.393	20.335	3.048***	−0.010
Observations	1,439,381	1,518,527	1,551,421			3,573,931	3,515,137	3,027,959		
Individuals	26,111	26,217	25,441			26,111	26,217	25,441		

Notes: 1. "Before" denotes the period that precedes the sickness period and starts on January 1999 and ends on June 2003 for the WAO group. It ends on September 2003 for the transitional WAO group, and on December 2003 for the WIA group. "After" denotes the period that succeeds the sickness period and starts on July 2004 and ends on December 2015 for the WAO group. It starts on October 2004 for the transitional WAO group, and on January 2006 for the WIA group. The transitional WAO and WIA groups are the treatment groups and consist of individuals who fell sick in the fourth quarter of 2003 and the first quarter of 2004, respectively, and are subjected to the reform. 2. Columns 1, 2, 3, 6, 7 and 8 present means of characteristics and outcomes of individuals in control and treatment groups before and after the sickness period. Columns 4, 5, 9 and 10 present the differences between treatment and control groups. A difference represents the estimated coefficient from the regression of the characteristic or outcome as the dependent variable, and the treatment indicator (whether member of the transitional WAO or the WIA group) as the explanatory variable. Standard error of the difference is clustered at the individual level. The reported F-statistic tests whether the estimated coefficients of all characteristics are jointly zero in a regression where treatment indicator is the dependent variable and all the individual's characteristics are included as explanatory variables. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively. 3. For background characteristics, a given mean represents the average of the individuals' given characteristic in the month these individuals fall sick. Therefore, the number of observations used to calculate the mean is the same as the presented number of individuals. For the outcome variables, the number of observations used in the before period is smaller than that in the after period since the period that precedes the sickness period is shorter than the period that succeeds it.

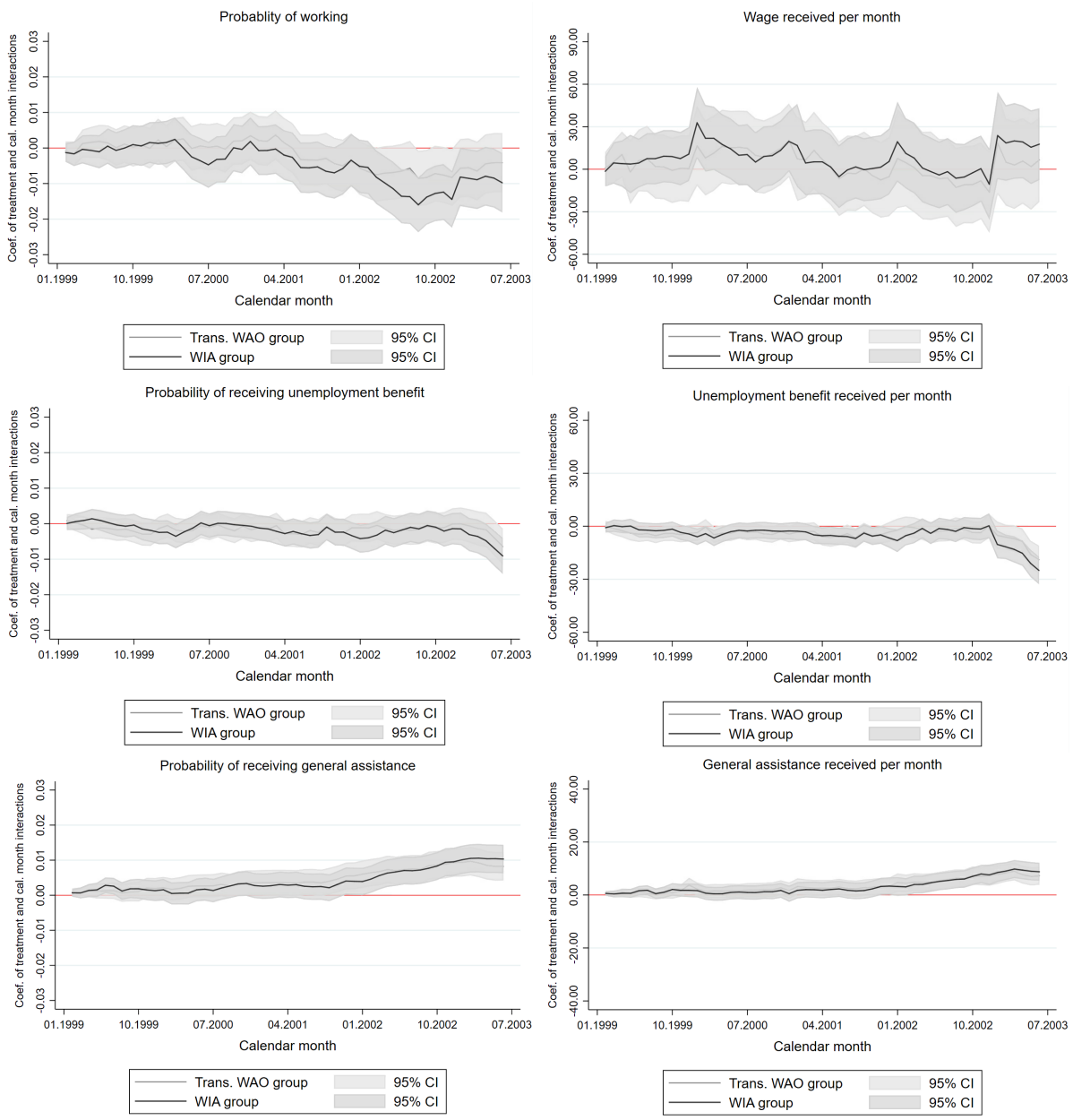

Figure 5a: Coefficient estimates of treatment and calendar month interactions in the pre-reform period from regressions for labor participation, benefit (UI and general assistance) receipt, labor income, and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on pre-reform data from January 1999 to June 2003. January 1999 is the base month for comparison. Labor participation, benefit receipt, labor income or benefit income is the outcome, and calendar month dummies, treatment and calendar month dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 4,199,526 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level. Test results based on the F-statistic of whether the coefficients of all interactions are jointly zero are as follows (p-values in parentheses). For labor participation, UI receipt, and general assistance receipt outcomes, the results are respectively 1.100 (0.287), 1.190 (0.164), 0.940 (0.595) for the transitional WAO group, and 1.560 (0.006), 1.320 (0.061), 1.600 (0.004) for the WIA group. For labor, UI, and general assistance income outcomes, the results are respectively 1.390 (0.030), 1.570 (0.005), 1.550 (0.006) for the transitional WAO group, and 1.330 (0.056), 1.990 (0.000), 2.070 (0.000) for the WIA group.

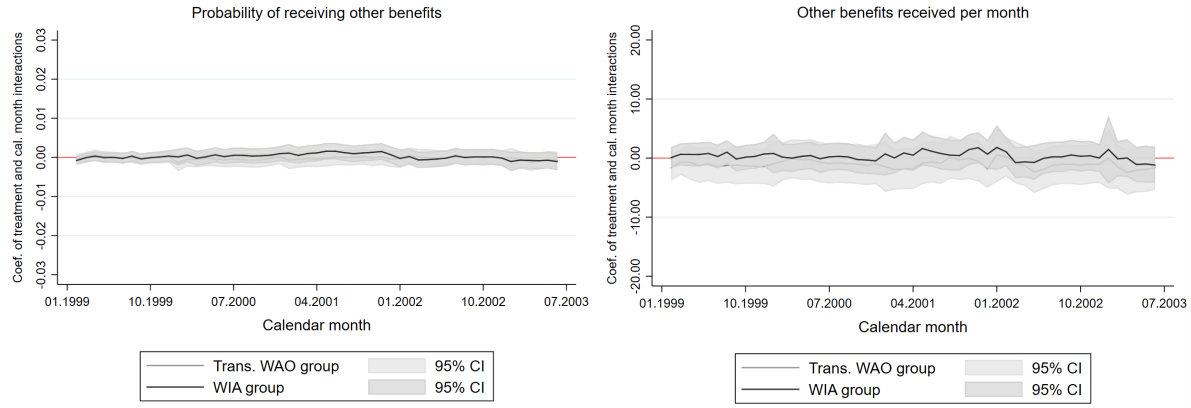

Figure 5b: Coefficient estimates of treatment and calendar month interactions in the pre-reform period from regressions for benefit (other benefits) receipt and income outcomes. Around each estimate is a 95 percent confidence interval. Regressions are based on pre-reform data from January 1999 to June 2003. January 1999 is the base month for comparison. Benefit receipt or income is the outcome, and calendar month dummies, treatment and calendar month dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 4,199,526 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level. Test results based on the F-statistic of whether the coefficients of all interactions are jointly zero are as follows (p-values in parentheses). For the other benefits receipt outcome, the result is 0.86 (0.747) for the transitional WAO group, and 1.220 (0.128) for the WIA group. For the other benefits income outcome, the result is 0.770 (0.885) for the transitional WAO group, and 0.96 (0.556) for the WIA group.

6 The effects of the disability reforms on work and benefit claiming behavior

6.1 Baseline effects of the reforms

Here we present the baseline DiD estimates of the effects of the transitional WAO and WIA reforms based on Equation (1). Table 2a presents the results for labor participation and benefit receipt, and Table 2b presents the results for income from labor and benefit programs. The estimated effects of the transitional WAO and the WIA reforms are always interpreted as the effects of the new rules of the transitional WAO and WIA regimes compared to the old rules of the WAO regime.

Table 2a shows that the WIA reform reduced the probability of receiving DI by 5.8 percentage points, on average, during the ten years after the reform has come into effect (where data is available on DI). The reduction of 5.8 percentage points corresponds to a reduction of 28.6 percent (0.058/0.203) in DI rewards among the individuals who fell sick during the third quarter of 2003 and the first quarter of 2004. This large effect of the WIA reform is in line with the finding of Van Sonsbeek and Gradus (2013) who showed that the reform reduced the number of DI rewards in the insured population of workers and unemployed by 34 percent and had a slightly diminishing effect over time.

As the WIA reform reduced DI receipt, it increased labor participation. The probability of working is 1.8 percentage points higher for the WIA group compared to the WAO group. This shows that the substantial decrease in DI receipt is not matched by a comparable increase in labor participation among the individuals in this group. It might be that DI recipients are often labor participants at the same time. We find, however, that nearly half of the decrease in DI receipt occurred among labor participants. Below we explore an alternative explanation that labor supply responses to the WIA reform occur at the intensive margin (labor participation and benefit receipt) instead of the extensive margin (labor and benefit income).

Individuals who could not access DI might have turned to benefits from other benefit programs. We find evidence that the WIA reform induced sick individuals to turn to UI. Compared to the comparison group, the WIA group increases its UI receipt, and the effect size is comparable to that of labor participation. The group, however, became less likely to receive general assistance. It might be that the income earned above the subsistence minimum due to work resumption or benefit substitution limits the access to general assistance. Another possibility is that there is less need for claiming general assistance on top of the DI if DI is lower than the applicable social minimum and is supplemented up to the social minimum by means of the Supplementary benefit (Section 2). Earlier studies in other countries also find evidence that tightening the eligibility criteria leads to more take-up of other benefits (Karlström et al., 2008; Staubli, 2011). Studies on earlier DI reforms in the Netherlands, however, show mixed results. Borghans et al. (2014) find that the disability reform that introduced reexaminations using stricter entitlement criteria in 1993 led to more benefit claims from other benefit programs, while Koning and van Vuuren (2010) and De Jong et al. (2011) find that the experience rating reform in 1998 and the Gatekeeper protocol reform in 2002 did not increase the probability of UI receipt.[14]

Table 2b shows the effects for (monthly) disability and unemployment benefits received and wages earned. Compared to the WAO regime, under the WIA regime individuals received €50.8 less disability benefits while they earned €55.7 more wages and received €22.8 more unemployment benefits, on average. They also received €4.9 less from general assistance. This

[14]The Gatekeeper protocol reform mandates firms to report sickness on time and clarifies firms' responsibilities of re-integrating the sick employees back to work.

shows that, due to the WIA reform, on average, individuals fully compensate lost DI benefits with higher earnings, but they also increase their UI claims so that, on balance, individuals receive a higher income (€22.8). This result is remarkable in view of the finding that the increase in labor participation due to the reform is rather modest. These effects are economically significant given the sizes of the DI, UI and general assistance received per month on average during the post-reform period (Table 1).

The effects for the transitional WAO group are considerably smaller and statistically less significant for all outcomes except for the benefits received from small benefit programs. This is expected since the WIA reform brought a much stricter benefit regime than the transitional WAO reform did.

6.2 Long-term effects of the reforms

We analyze the long-term effects of the transitional WAO and WIA reforms in Figures 6a and 6b. In a given subfigure of a figure, we present the coefficient estimates of treatment and year dummy interactions over the post-treatment period of ten years based on the regression given by Equation (2). In line with the exploratory graphical analysis in Figures 3a and 3b, the WIA reform shows structural effects on DI receipt, income from DI, labor participation, and earnings. During the entire post-treatment period, the effect of the reform on these outcomes are sizable and statistically significant. The effect of the reform on UI is large at the onset of the reform but fades out in about seven years after the reform is introduced. This a consequence of the UI being temporary. Similar patterns are observed for general assistance and other benefits. The impact of the transitional WAO reform on labor participation and benefit use is short-lived. For example, as discussed in Section Section 4, the reform has a significant effect on DI receipt only during the first two years of the reform. After the first two years, all participants in the old WAO scheme who were younger than 45 years of age were re-examined based on the eligibility criteria of the transitional WAO scheme. Therefore, the differences between old and transitional WAO scheme diminished for the younger participants.

6.3 Heterogeneous effects of the reforms

Here we analyze whether the disability reforms affected individuals with different background and labor market characteristics differently. We consider differential effects for gender, age, work status, contract type, and occupation type. The age groups distinguish between individuals who are younger than 40, between 40 and 44, between 45 and 49, between 50 and 54, between 55 and 59, and older than 59. The work status and contract type groups distinguish between individuals who hold a permanent contract, a temporary contract or a contract through a temporary work agency, and who are unemployed. Occupation groups distinguish between white-collar workers, blue-collar workers, and workers employed through temporary work agencies. Age, work status, and contract and occupation types are those at the time individuals fall sick. Tables 3a through 6b present the results for labor participation, benefit receipt, and labor and benefit income based on the regression given by Equation (1) on these subgroups. We focus on the effects of the WIA reform as the effects of the transitional WAO reform are often insignificant.

Tables 3a and 3b show that men are less likely to receive DI and they receive a smaller DI compared to women. Based on the Wald test, the difference between the coefficient estimates of the two groups is statistically significant at the 10 percent level for both DI receipt and income. This result is expected. Men earn, on average, higher wages than women. In the Dutch disability scheme, a higher pre-sickness wage means a higher probability of getting a partial DI. This is because the pre-sickness wage is compared to a (lower) fictitious new wage

that can be earned given one's disability. A higher pre-sickness wage implies more alternative work opportunities with lower fictitious new wages (Section 2).

Tables 4a and 4b show notable age effects. First, compared to the younger and older age groups, the 50-54 age group is affected the most in terms of decreased DI receipt and income. This seems to explain why the same group increases its labor participation the most and earns the highest wages. To a certain extent these findings might reflect a composition effect. Individuals of the 50-54 age group more often hold a permanent work contract (68 percent versus 49 and 61 percent for the youngest and oldest age groups, respectively) which might make it more difficult to access DI. For example, it might be easier to find potential jobs these individuals can still perform leading to a smaller disability grade (Section 2).

Second, compared to the younger age groups, the oldest age group is much more likely to rely on disability and unemployment benefits in terms of both benefit receipt and income. For example, the UI received by this group is almost seven times larger than that received by the youngest age group. This group is also much less likely to increase labor participation and earnings. In 2013, the government has put into effect the law raising the legal retirement age. The results show that sick individuals close to the legal retirement age very much struggle to respond to the work incentives of the reforms. This suggests that a higher legal retirement age is likely to limit the remaining work capacity further and increase benefit claims or amounts among the future population of disabled workers reaching the legal retirement age. Labor participation of this group can therefore be expected to become important in the coming decades, and even more so as the size of this group is increasing since the law raising the legal retirement age has come into effect. In fact, the number of benefit claims in the WIA increased remarkably from 37,140 in 2013 to 43,419 in 2018 (Berendsen et al., 2019).

The effect of the WIA reform shows substantial heterogeneity with respect to work status. First, it decreases the use of DI among the unemployed individuals more than it does among those who hold a temporary or permanent contract. It might be that the unemployed are more often sick, without necessarily having severe health problems, and therefore are more likely to be affected by a stricter DI regime. Furthermore, the reform decreases labor participation among the unemployed but increases it among those with permanent or temporary work contracts. The results based on the income outcomes show that unemployed individuals receive much larger amounts of UI due to loss of income from DI and wages. It might be that the work resumption incentives brought by the WIA reform lead employers to reintegrate their employees into their jobs while they prove ineffective if there is no employer. On the employees' side, the incentives to resume working may be smaller for the unemployed individuals insured under the WIA, compared to those insured under the WAO, since they spend an additional year in the sickness scheme before they become eligible to apply for DI. A longer unemployment spell may lead to more human capital loss or have a stronger "scarring effect" for the sick unemployed individuals, and decrease their prospects of finding a job (Arulampalam, 2001; Arulampalam et al., 2001). These results suggest that the unemployed are much more limited in their access to DI and lack the incentives the group of individuals with permanent or temporary contracts have to resume working.

A second notable result follows from comparing the results for workers with permanent and temporary contracts. Table 5a shows that, due to the reform, individuals with permanent contracts show a slightly larger increase in labor participation compared to those with temporary contracts. This is expected because these individuals have an employer to return to after recovering from sickness within two years. However, Table 5b shows that these individuals are far better in coping with the loss of DI as they increase their wages by a much larger amount and rely less on UI. Possibly for the same reason, resuming work with their own employer gives them more flexibility in increasing their number of work hours to the former level compared

to the individuals with temporary contracts who usually have to resume working with a new employer.

Finally, Tables 6a and 6b show reform effects by occupation type.[15] Blue-collar workers reduce their DI receipt more than those in other occupations. They increase labor participation but less than those in other occupations. They also increase earnings but rely on UI to fully compensate for lost DI. White-collar workers, on the other hand, overcompensate lost DI with earnings by a large margin. Those who are employed through temporary work agencies stand out with their ability to increase labor participation, but they fail to compensate lost DI with earnings and unemployment benefits.

Both supply and demand side factors can explain the large differences across occupation groups. Blue-collar workers often occupy physically demanding jobs and may find it more difficult to resume working after a health shock. They may also have limited job alternatives at their employer. This is consistent with Ravesteijn et al. (2018) who show that, in Germany, blue-collar workers report worse health than white-collar workers due to physical strain and low job control in their occupations. On the other hand, white-collar workers can increase earnings more in response to the reform. They earn higher hourly wages and have better opportunities to find jobs that fit their disability possibly because the jobs they do are technologically more compatible with part-time work or flexible hours (Hutchens and Grace-Martin, 2006) or because they more often work at large firms which are more successful in reintegrating disabled workers into their jobs (Van Sonsbeek and Gradus, 2013). Workers employed through temporary work agencies perform better in resuming work after becoming disabled most likely because for them it is easy to find temporary jobs to utilize remaining work capacity. However, they perform worse than those in other occupations to increase earnings. These people lose their employer as soon as they fall sick. Therefore, it might be that for the ones with an employer the increase in earnings mainly comes from (earlier) return to their own job. It might also be that they are confined to lower quality jobs, especially after a history of sickness. Overall, these results show that how sick people respond to stricter criteria for DI depends, by large, on the type of their occupation.

[15] Individuals in the sickness data take one of 69 occupation titles according to the industrial classification of all economic activities by UWV. We classify the 69 occupation titles into major occupation groups as follows. White-collar are people who work at public services, business services, financial institutions, cultural institutions, and wholesale and retail trade. Blue-collar are people who work in construction, transportation, agriculture, forestry, metal industry, food industry, among many other elementary occupations. These blue- and white-collar definitions are consistent with, for example, Ravesteijn et al. (2018) who distinguish between blue- and white-collar workers to study the impact of occupation type on health in Germany. People employed through temporary work agencies, or those who work for these agencies, define our third group. Both blue- and white-collar workers can be employed through these agencies. Although these workers can be hired by any company, the temporary work agency is their formal employer that pays social security contributions to UWV that administers sickness and disability insurance. As soon as these people fall sick, their contract with the temporary work agency is terminated and responsibility for sickness benefit payment and reintegration is taken over by UWV. According to these definitions, 32.5, 57.5 and 9.2 percent of the study sample represent white-collar workers, blue-collar workers, and workers employed through temporary work agencies, respectively. The remaining accounts for workers with a missing occupation title.

Table 2a: Linear model explaining the effects of disability reforms on the probabilities of labor participation and benefit receipt

	DI (and possibly UI) receipt	Labor part.	UI (no DI) receipt	General assistance receipt	Other benefits receipt
$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.008^{**}	0.007^{**}	0.003^{**}	-0.003^{*}	0.001^{***}
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.058^{***}	0.018^{***}	0.014^{***}	-0.004^{***}	-0.005^{***}
	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)
Post	0.203^{***}	-0.368^{***}	0.010^{***}	0.009^{***}	0.003^{***}
	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
R-squared (within)	0.124	0.182	0.012	0.006	0.002
Observations	14,626,356				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Table 2b: Linear model explaining the effects of disability reforms on the amounts of wage and benefits received per month

	DI (and possibly UI) received per month	Wages earned per month	UI (excl. of DI) received per month	General assistance received per month	Other benefits received per month
$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-5.528	19.284	5.319^{***}	-2.841^{*}	2.262^{*}
	(4.300)	(12.802)	(1.882)	(1.487)	(1.225)
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-50.813^{***}	55.659^{***}	22.824^{***}	-4.898^{***}	0.248
	(4.413)	(13.332)	(1.952)	(1.550)	(1.230)
Post	282.124^{***}	-389.573^{***}	17.368^{***}	19.042^{***}	11.393^{***}
	(3.411)	(11.262)	(1.838)	(1.371)	(2.721)
R-squared (within)	0.081	0.036	0.010	0.007	0.000
Observations	14,626,356				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

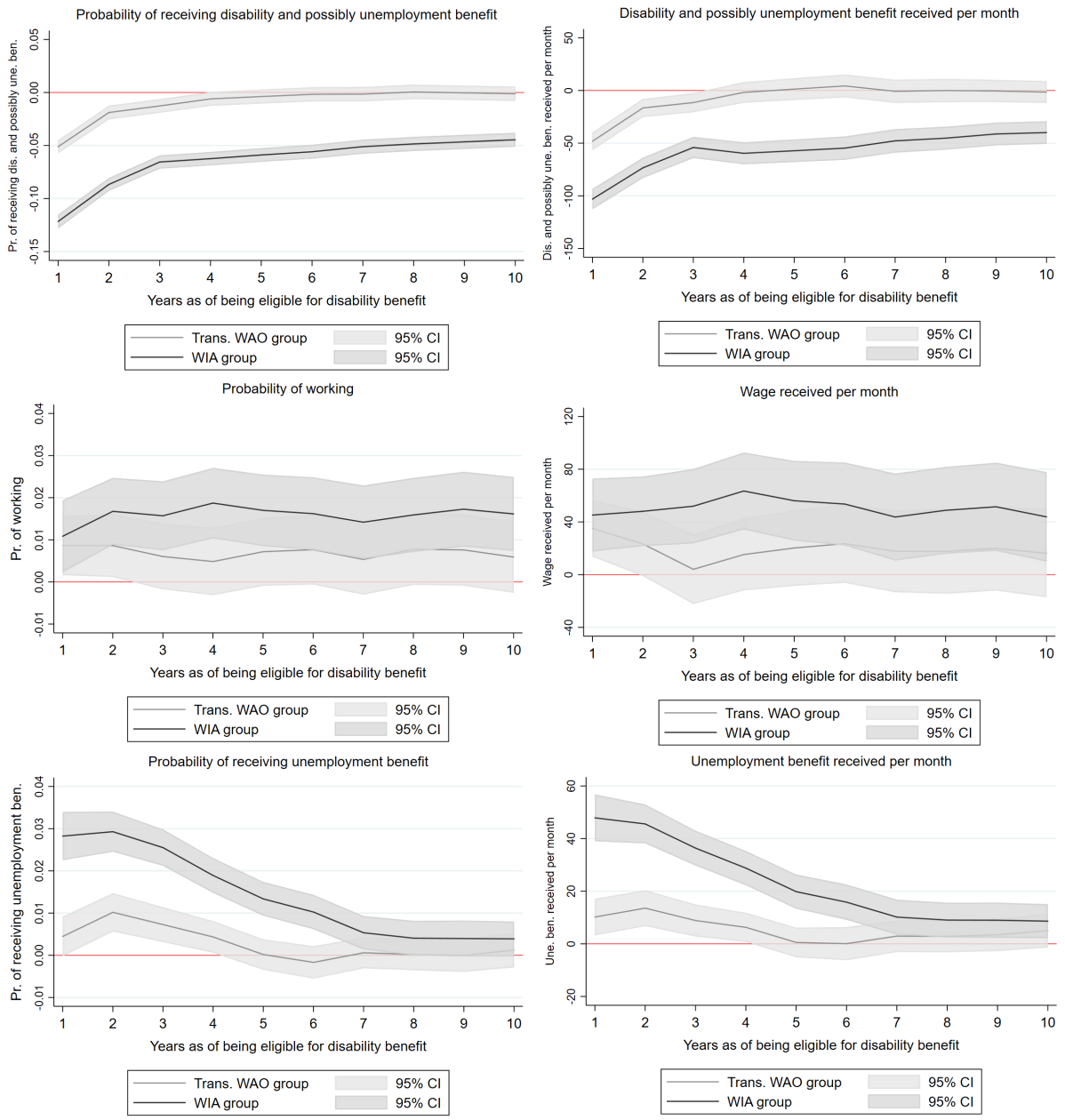

Figure 6a: Coefficient estimates of treatment and annual dummy interactions for control and treatment groups during the reform period from regressions for labor participation, benefit (DI and UI) receipt, labor income and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

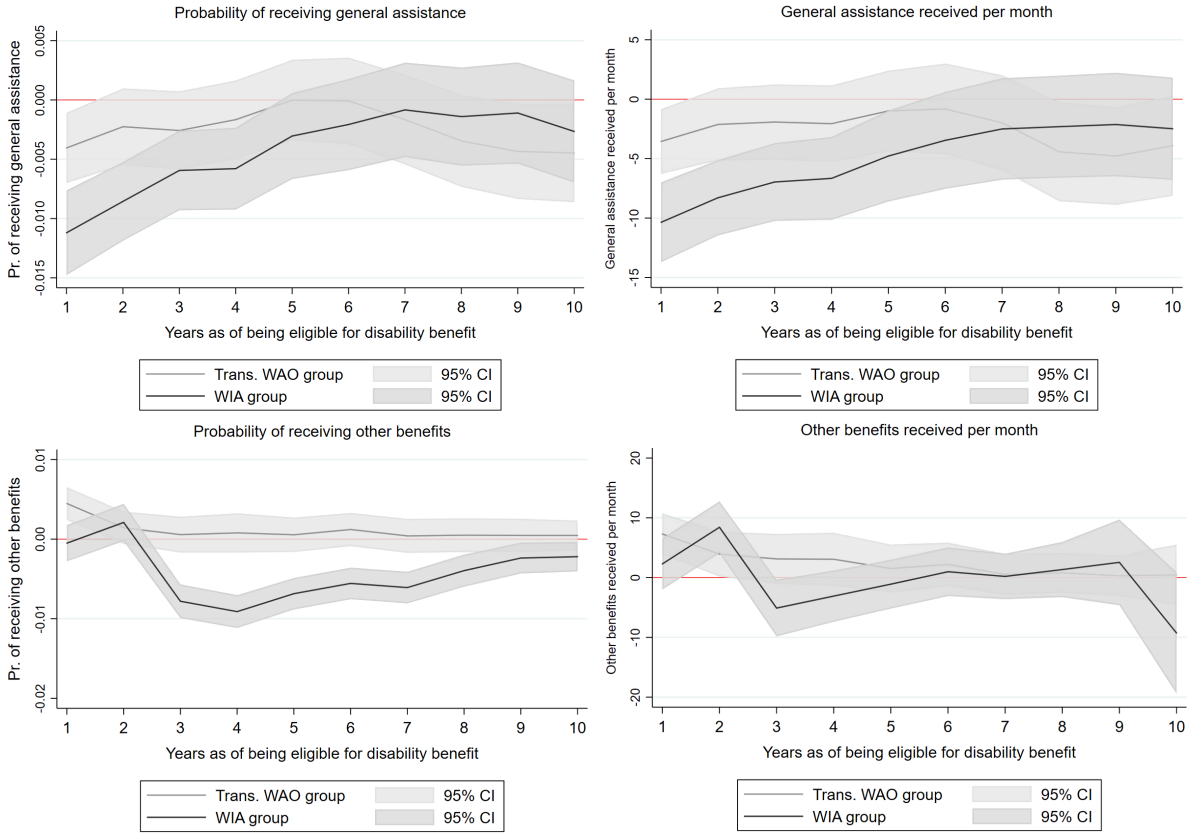

Figure 6b: Coefficient estimates of treatment and annual dummy interactions for control and treatment groups during the reform period from regressions for benefit (general assistance and other benefits) receipt and income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Benefit receipt or income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

Table 3a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by gender groups

		Female	Male
DI (and possibly UI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.006	-0.009^{**}
		(0.004)	(0.004)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.049^{***}	-0.066^{***}
		(0.004)	(0.004)
Labor participation	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.009^{*}	0.006
		(0.005)	(0.005)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.019^{***}	0.017^{***}
		(0.005)	(0.005)
UI (no DI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.004^{*}	0.002
		(0.002)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.013^{***}	0.015^{***}
		(0.002)	(0.002)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.002	-0.004^{**}
		(0.003)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.005^{**}	-0.003^{*}
		(0.003)	(0.002)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	0.001
		(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{***}	-0.007^{***}
		(0.001)	(0.001)
Observations		6,429,372	8,196,984
Individuals		34,186	43,583

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for men and women are as follows. For the transitional WAO reform, they are 0.596, 0.671, 0.480, 0.579, and 1.000, respectively for DI receipt, labor participation, UI receipt, general assistance receipt, and other benefit receipt outcomes. For the WIA reform, they are 0.003, 0.777, 0.480, 0.579, 0.005 for the respective outcomes.

Table 3b: Linear model explaining the effects of disability reforms on labor and benefit income by gender groups

		Female	Male
DI (and possibly UI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-9.903^{*}	-2.629
		(5.551)	(6.288)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-42.663^{***}	-57.375^{***}
		(5.720)	(6.454)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	21.049	20.141
		(14.452)	(19.657)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	45.298^{***}	64.482^{***}
		(15.289)	(20.395)
UI (excl. of DI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	5.544^{**}	5.125^{*}
		(2.277)	(2.844)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	15.630^{***}	28.468^{***}
		(2.358)	(2.953)
General assistance per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-1.970	-3.536^{**}
		(2.637)	(1.662)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-6.059^{**}	-3.996^{**}
		(2.738)	(1.739)
Other benefits per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	1.071	3.138
		(1.333)	(1.918)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-1.547	1.642
		(1.312)	(1.939)
Observations		6,429,372	8,196,984
Individuals		34,186	43,583

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time, and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for men and women are as follows. For the transitional WAO reform, they are 0.386, 0.970, 0.908, 0.615, and 0.376, respectively for DI, labor, UI, general assistance, and other benefits income outcomes. For the WIA reform, they are 0.088, 0.452, 0.001, 0.525, 0.173 for the respective income outcomes.

Table 4a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by the age when individuals fell sick

		Fell sick before age 40	Fell sick between ages 40 and 44	Fell sick between ages 45 and 49	Fell sick between ages 50 and 54	Fell sick between ages 55 and 59	Fell sick after age 59
DI (and possibly UI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.003	-0.015^{*}	-0.002	-0.029^{***}	-0.004	-0.011
		(0.004)	(0.008)	(0.009)	(0.010)	(0.008)	(0.011)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.046^{***}	-0.067^{***}	-0.081^{***}	-0.091^{***}	-0.051^{***}	-0.031^{***}
		(0.004)	(0.008)	(0.009)	(0.010)	(0.007)	(0.009)
Labor participation	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.008	0.014^{*}	0.010	0.030^{***}	0.011	0.002
		(0.005)	(0.009)	(0.009)	(0.009)	(0.009)	(0.022)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.003^{***}	0.021^{**}	0.023^{**}	0.043^{***}	0.013	0.004
		(0.005)	(0.009)	(0.010)	(0.010)	(0.009)	(0.021)
UI (no DI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	-0.000	0.004	0.004	0.010^{**}	0.024^{*}
		(0.001)	(0.003)	(0.004)	(0.004)	(0.005)	(0.014)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.007^{***}	0.013^{***}	0.016^{***}	0.029^{***}	0.030^{***}	0.033^{**}
		(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.014)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.004	-0.004	0.001	-0.002	0.000	-0.006
		(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.007)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.004	-0.006	-0.002	-0.004	0.002	-0.006
		(0.002)	(0.004)	(0.004)	(0.004)	(0.003)	(0.007)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	-0.001	0.002	0.000	0.003	-0.002
		(0.001)	(0.002)	(0.003)	(0.003)	(0.003)	(0.005)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{***}	-0.008^{***}	-0.006^{***}	-0.010^{***}	-0.003	-0.002
		(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)
Observations		7,099,392	2,084,652	1,864,368	1,782,324	1,406,172	389,448
Individuals		37,724	11,087	9,915	9,481	7,486	2,076

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time, and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Table 4b: Linear model explaining the effects of disability reforms on labor and benefit income by the age when individuals fell sick

		Fell sick before age 40	Fell sick between ages 40 and 44	Fell sick between ages 45 and 49	Fell sick between ages 50 and 54	Fell sick between ages 55 and 59	Fell sick after age 59
DI (and possibly UI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−0.533	−13.981	2.763	−46.739***	9.442	−17.428
		(5.080)	(11.689)	(14.824)	(16.094)	(13.364)	(12.893)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−37.139***	−48.792***	−81.611***	−103.513***	−44.784***	−22.614*
		(5.266)	(12.263)	(15.017)	(16.631)	(12.853)	(12.749)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	31.518**	50.331	32.396	108.105***	−3.782	−27.985
		(15.788)	(31.254)	(36.250)	(34.607)	(44.997)	(80.652)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	95.659***	83.539**	96.918***	157.964***	58.700	−31.836
		(16.581)	(32.711)	(36.458)	(37.188)	(43.227)	(78.772)
UI (excl. of DI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	2.366	3.962	8.858	2.125	14.837*	47.416*
		(1.919)	(4.704)	(5.596)	(6.360)	(8.689)	(24.411)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	10.981***	22.582***	21.292***	41.984***	56.365***	69.712***
		(1.995)	(4.822)	(5.776)	(6.730)	(8.620)	(23.916)
General assistance per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−4.123*	−4.028	2.311	−1.908	1.315	−2.656
		(2.390)	(4.021)	(3.966)	(3.412)	(2.507)	(4.329)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−5.174**	−5.769	−2.447	−4.293	2.380	−5.393
		(2.520)	(4.227)	(4.087)	(3.604)	(2.434)	(4.470)
Other benefits per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	1.216	2.917	2.942	0.193	11.794*	−18.181
		(0.927)	(2.635)	(4.217)	(5.370)	(6.243)	(12.476)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−0.249	2.705	−2.649	−3.681	6.038	−2.834
		(0.987)	(2.606)	(3.641)	(6.050)	(5.817)	(10.497)
Observations		7,099,392	2,084,652	1,864,368	1,782,324	1,406,172	389,448
Individuals		37,724	11,087	9,915	9,481	7,486	2,076

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Table 5a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by work status

		Permanent contract	Temporary contract	Unemployed
DI (and possibly UI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.009^{**}	-0.016^{**}	-0.004
		(0.004)	(0.008)	(0.008)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.058^{***}	-0.055^{***}	-0.068^{***}
		(0.003)	(0.008)	(0.008)
Labor participation	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.008^{*}	0.019^{**}	-0.000
		(0.004)	(0.008)	(0.008)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.032^{***}	0.026^{***}	-0.027^{***}
		(0.005)	(0.009)	(0.008)
UI (no DI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	0.007^{**}	0.007^{*}
		(0.001)	(0.003)	(0.004)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.005^{***}	0.023^{***}	0.043^{***}
		(0.001)	(0.003)	(0.004)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.001	-0.009^{*}	-0.006
		(0.001)	(0.005)	(0.004)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{*}	-0.014^{***}	-0.003
		(0.001)	(0.005)	(0.005)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	-0.002	0.003
		(0.001)	(0.002)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.006^{***}	0.002	-0.009^{***}
		(0.001)	(0.002)	(0.002)
Observations		8,518,944	2,534,856	2,945,844
Individuals		45,312	13,445	15,682

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Work status indicates individuals who hold a permanent contract, temporary contract or a contract through a temporary work agency, or are unemployed at the time of falling sick. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals who hold a temporary work contract and those who hold a permanent work contract are as follows. For the transitional WAO reform, they are 0.434, 0.219, 0.058, 0.117, and 0.180, respectively for the DI receipt, labor participation, UI receipt, general assistance receipt, and other benefits receipt outcomes. For the WIA reform, they are 0.725, 0.560, 0.000, 0.031, 0.000 for the respective outcomes. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals who are unemployed and those who hold a permanent work contract are as follows. For the transitional WAO reform, they are 0.576, 0.371, 0.146, 0.225, and 0.371, respectively for the DI receipt, labor participation, UI receipt, general assistance receipt, and other benefits receipt outcomes. For the WIA reform, they are 0.242, 0.000, 0.000, 1.000, 0.180 for the respective outcomes.

Table 5b: Linear model explaining the effects of disability reforms on labor and benefit income by work status

		Permanent contract	Temporary contract	Unemployed
DI (and possibly UI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−4.681	−12.497	−9.387
		(5.231)	(10.595)	(11.194)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−55.753***	−39.572***	−51.530***
		(5.292)	(11.544)	(11.454)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	30.114	45.490*	−9.080
		(18.607)	(25.687)	(23.914)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	125.315***	61.299**	−119.296***
		(19.240)	(28.120)	(23.816)
UI (excl. of DI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	2.518	10.605***	18.325***
		(2.007)	(3.937)	(6.169)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	8.604***	32.785***	70.116***
		(2.140)	(4.301)	(6.107)
General assistance per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−1.532	−9.038*	−5.472
		(1.270)	(4.723)	(4.468)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−2.974**	−15.257***	−3.515
		(1.345)	(5.062)	(4.631)
Other benefits per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	3.493**	−0.336	2.592
		(1.753)	(2.254)	(2.677)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−3.611**	6.693**	6.562***
		(1.733)	(2.818)	(2.450)
Observations		8,518,944	2,534,856	2,945,844
Individuals		45,312	13,445	15,682

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Work status indicates individuals who hold a permanent contract, temporary contract or a contract through a temporary work agency, or are unemployed at the time of falling sick. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals who hold a temporary work contract and those who hold a permanent work contract are as follows. For the transitional WAO reform, they are 0.508, 0.628, 0.067, 0.125, and 0.180, respectively for the DI, labor, UI, general assistance, and other benefits income outcomes. For the WIA reform, they are 0.203, 0.060, 0.000, 0.019, 0.002 for the respective income outcomes. P-values based on the Wald statistic for the equality of the estimated reform effects for the individuals who are unemployed and those who hold a permanent work contract are as follows. For the transitional WAO reform, they are 0.703, 0.196, 0.015, 0.396, and 0.778, respectively for the DI, labor, UI, general assistance, and other benefits income outcomes. For the WIA reform, they are 0.738, 0.000, 0.000, 0.911, 0.001 for the respective income outcomes.

Table 6a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by occupation type

		Blue-collar	White-collar	Temporary work agency
DI (and possibly UI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.007^{*}	-0.011^{**}	-0.001
		(0.004)	(0.005)	(0.010)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.063^{***}	-0.051^{***}	-0.047^{***}
		(0.004)	(0.005)	(0.010)
Labor participation	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.009^{*}	0.001	0.026^{**}
		(0.005)	(0.006)	(0.011)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.017^{***}	0.020^{***}	0.030^{**}
		(0.005)	(0.006)	(0.012)
UI (no DI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.001	0.004^{*}	0.004
		(0.002)	(0.002)	(0.005)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.014^{***}	0.011^{***}	0.020^{***}
		(0.002)	(0.002)	(0.005)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.002	-0.002	-0.012^{*}
		(0.002)	(0.002)	(0.007)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003	-0.003	-0.007
		(0.002)	(0.002)	(0.007)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.002^{*}	-0.001	-0.001
		(0.001)	(0.001)	(0.002)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.006^{***}	-0.006^{***}	0.000
		(0.001)	(0.001)	(0.002)
Observations		8,402,292	4,749,336	1,352,448
Individuals		44,687	25,267	7,171

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. White-collar are people who work at public services, business services, financial institutions, cultural institutions, and wholesale and retail trade. Blue-collar are people who work in construction, transportation, agriculture, forestry, metal industry, food industry, among many other elementary occupations. Temporary work agency workers are those employed through temporary work agencies or those who work for these agencies. Both blue- and white-collar workers can be employed through these agencies. Occupation type is one at the time of falling sick. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Table 6b: Linear model explaining the effects of disability reforms on labor and benefit income by occupation type

		Blue-collar	White-collar	Temporary work agency
DI (and possibly UI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−4.268	−11.729	2.952
		(5.668)	(7.766)	(13.043)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−51.297***	−51.984***	−42.078***
		(5.793)	(7.944)	(13.810)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	24.296	12.510	30.102
		(15.558)	(27.335)	(27.617)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	35.601**	111.969***	32.084
		(15.914)	(28.352)	(30.725)
UI (excl. of DI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	3.901	7.139**	6.702
		(2.403)	(3.471)	(5.931)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	20.912***	21.571***	30.253***
		(2.453)	(3.646)	(6.455)
General assistance per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	−1.881	−2.479	−11.356
		(1.965)	(2.027)	(6.911)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	−3.519*	−3.002	−12.439*
		(2.058)	(2.162)	(7.420)
Other benefits per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	2.193	2.890	−0.168
		(1.367)	(2.866)	(1.922)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	1.265	−2.782	1.188
		(1.349)	(2.856)	(2.264)
Observations		8,402,292	4,749,336	1,352,448
Individuals		44,687	25,267	7,171

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. White-collar are people who work at public services, business services, financial institutions, cultural institutions, and wholesale and retail trade. Blue-collar are people who work in construction, transportation, agriculture, forestry, metal industry, food industry, among many other elementary occupations. Temporary work agency workers are those employed through temporary work agencies or those who work for these agencies. Both blue- and white-collar workers can be employed through these agencies. Occupation type is one at the time of falling sick. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

7 Conclusion

In the last decades DI programs have grown in many western countries. Governments implement social security reforms to reduce enrollment in the DI program and the cost of DI. We evaluated a major DI reform that introduced strong incentives for work resumption, not only in the disability scheme but also in the sickness scheme that precedes it. Using unique administrative data and difference-in-difference analysis on individuals who fall sick before and after the reform, we analyzed the effect of the reform on labor participation decisions and use of benefits from alternative benefit programs as well as on earnings and income recieved from these programs.

The new benefit regime reduced DI receipt by 5.8 percentage points (28.6 percent) among sick-listed workers, which confirms earlier findings by Van Sonsbeek and Gradus (2013) who estimated the new benefit regime to reduce DI receipt in the insured population of workers and unemployed by 34 percent after the reform compared to the pre-reform period, and found a slightly diminishing effect of the reform over time. We extend the evidence on DI use in several respects and provide a broad picture of the many different effects of the DI reform. First, we show that the reform increased labor participation by 1.8 percentage points, about a third of the decrease in DI receipt. This points to a modest labour participation effect. The reform also increased UI receipt by 1.4 percentage points, which equals a fourth of the decrease in DI receipt. The 1.8 percentage points labour participation effect is smaller than the 2.9 percentage points effect found in Borghans et al. (2014), and the 6.7 percentage points effect found in Garcia-Mandicó et al. (2020), who analyzed the effects of reexamination of existing DI recipients. Comparing these results is difficult, however, because of the use of different study populations. The populations in Borghans et al. and Garcia-Mandicó et al. consist of long-term DI beneficiaries, among which only few were already working during the pre-reform period (before their reexamination), whereas our population consists of sick-listed workers of which the majority were working before sickness and most recover before they reach DI eligibility.

Second, we show substantial effects at the intensive margin. The reform cohort lost on average €51 per month (18 percent) in DI benefits, but increased monthly wages by €56 and income from UI by €23. This means that individuals fully compensate lost DI by increasing earnings (110 percent) and they overcompensate with additional income from UI (45 percent). Borghans et al. (2014) and Garcia-Mandicó et al. (2020) show a substantially lower earnings rebound of 62 and 64 percent, respectively, whereas the rate of social support substitution is comparable to the 30 percent found by Borghans et al.. Apparently, the fact that the new system intervenes early in the sickness period (the most important changes take place within two years after the first day of sickness) allows for a stronger earnings rebound than a late intervention by means of reexamination of people that are already on DI benefits. This is a confirmation of one of the most important policy lessons learned in the Netherlands after subsequent reforms of the DI schemes: the earlier an intervention is made after sickness, the better the chances of (partial) recovery and work resumption.

Third, we show that the impact of the reform had been persistent during the ten years of the study period. This is an important conclusion from a policy perspective as well because earlier reforms failed to remain effective in the long run. This in particular holds for the major DI reform that took place in 1993 in which entitlement criteria were tightened. Whereas the effect of this reform on reexaminations, as described in Borghans et al. (2014), sustained, the effect of the same reform to new cohorts was short-lived. After an initial sharp drop in annual inflow, the amount of new DI receipts was back to its pre-reform level within five years because of a gradual relaxation of the entrance examinations.

Fourth, the impact of the reform is substantially heterogenous with respect to work status and age. Individuals who are on a permanent contract are better able to compensate loss of DI

benefits with higher earnings than those with a temporary contract. This is evidence of labor market segmentation, which in fact has grown stronger over time in the Netherlands where differences between a relatively well-off and low-risk group of insiders contrast with a group of less wealthy, less healthy and higher risk group of outsiders. Even more vulnerable than the workers with a temporary contract are the unemployed, who lack the incentives of the group of individuals with permanent or temporary contracts to resume working. The unemployed are not able at all to compensate lost disability benefits by either wages or alternative benefits. This suggests that unemployed individuals face additional limitations in their access to the labor market when they fall sick.

When considering the reform effects by age, older individuals appear as a very vulnerable group. They are less able to compensate the decrease in DI receipt or income with higher labor participation or wages, whereas they more often rely on UI. This confirms the weak labour market position and low labor mobility of older workers in the Netherlands (Visser et al., 2018). Furthermore, as the statutory pension age is being raised, later retirement is likely to worsen the labor market position of disabled individuals approaching retirement age who struggle the most to utilize their remaining work capacity.

Furthermore, how sick people respond to stricter criteria for DI depends, by large, on occupation type. On average, white-collar workers manage to substantially overcompensate lost DI with earnings while blue-collar workers must rely on UI to compensate because they struggle more to cope with their disability. Those who are employed through temporary work agencies fail to compensate. When these workers fall sick, unlike workers with a permanent contract, they do not have an employer who is responsible to compensate the lost wages and is incentivized to facilitate work resumption. The low wages they earn contribute to their income insecurity and make them particularly vulnerable to stricter rules for DI.

These heterogeneous effects raise inequality concerns and call for attention to vulnerable labour market groups when such a far-reaching reform is implemented, such as blue-collar workers with physically demanding jobs, workers who rely on temporary contracts for a long time, the unemployed, and older workers with health conditions. Like in the WIA reform, employers are supposed to play an important role in facilitating work resumption, but they are not eager to employ workers from vulnerable groups, especially when there are strong financial incentives for employers to re-integrate if their workers become sick or disabled. Therefore, comprehensive reforms should pay special attention to protecting vulnerable groups from welfare loss.

In fact, the share of the vulnerable groups in the Dutch labor force is increasing. The number of workers with temporary contracts increased from about 1 to 2 million between 2003 and 2018. Two-third of this group preferred a permanent contract in 2016 (Salverda, 2020). Furthermore, the raising of the statutory retirement age since 2013 is increasing the share of older people in the labor force. These compositional changes of the labor force suggest that future cohorts of DI applicants will include more people from vulnerable groups. Therefore, it becomes increasingly important to understand how DI reforms such as the one studied here affect benefit claiming and substitution across different segments of the labor force.

Acknowledgments

This research is supported by the Network for studies on Pensions, Aging and Retirement (Netspar) under grant number LMVP 2014.03. Its contents are the responsibility of the authors and do not necessarily represent the official views of Netspar. We thank UWV, and in particular Lucien Rondagh and Roel Ydema, for providing the sickness data. We thank Jennifer Alonso-García, Meltem Daysal, Jochem de Bresser, Pilar García-Gómez, Marike Knoef, Pierre Koning,

Marina Krauß, Martin Salm, Jan van Ours, Arthur van Soest, and the conference participants at the Netspar Pension Day 2018, Netspar International Pension Workshop 2019, and 6th Workshop of the DGGÖ Health Econometrics Working Group 2019 for their helpful comments.

References

Arulampalam, W., 2001. Is unemployment really scarring? Effects of unemployment experiences on wages. The Economic Journal 111 (475), F585–606.

Arulampalam, W., Gregg, P., Gregory, M., 2001. Introduction: unemployment scarring. The Economic Journal 111 (475), F577–584.

Autor, D. H., Duggan, M., Greenberg, K., Lyle, D. S., 2016. The impact of disability benefits on labor supply: Evidence from the VA's disability compensation program. American Economic Journal: Applied Economics 8 (3), 31–68.

Autor, D. H., Duggan, M. G., 2003. The rise in the disability rolls and the decline in unemployment. The Quarterly Journal of Economics 118 (1), 157–206.

Berendsen, E., van Deursen, C., van Sonsbeek, J.-M., 2019. Stijging WIA-instroom nog geen reden tot zorg. Economisch Statistische Berichten 104 (4773), 230–232.

Borghans, L., Gielen, A. C., Luttmer, E. F. P., 2014. Social support substitution and the earnings rebound: evidence from a regression discontinuity in disability insurance reform. American Economic Journal: Economic Policy 6 (4), 34–70.

Burkhauser, R., Marc, D., McVicar, D., Wilkins, R., 2014. Disability benefit growth and disability reform in the US: lessons from other OECD nations. IZA Journal of Labor Policy 3 (4), 1–30.

Campolieti, M., 2004. Disability insurance benefits and labor supply: some additional evidence. Journal of Labor Economics 22 (4), 863–889.

Campolieti, M., Riddell, C., 2012. Disability policy and the labor market: evidence from a natural experiment in Canada, 1998-2006. Journal of Public Economics 96 (3-4), 306–316.

Chen, S., van der Klaauw, W., 2008. The work disincentive effects of the disability insurance program in the 1990s. Journal of Econometrics 142 (2), 757–784.

Dahl, G. B., Gielen, A. C., 2018. Intergenerational spillovers in disability insurance. NBER Working Paper No. 24296.

De Groot, N., Koning, P., 2016. Assessing the effects of disability insurance experience rating. The case of the Netherlands. Labour Economics 41, 304–317.

De Jong, P., Lindeboom, M., 2004. Privatisation of sickness insurance: evidence from the Netherlands. Swedish Economic Policy Review 11, 121–144.

De Jong, P., Lindeboom, M., van der Klaauw, B., 2011. Screening disability insurance applications. Journal of the European Economic Association 9 (1), 106–129.

Deshpande, M., 2016. The effect of disability payments on household earnings and income: Evidence from the SSI children's program. Review of Economics and Statistics 98 (4), 638–654.

Deuchert, E., Eugster, B., 2019. Income and substitution effects of a disability insurance reform. Journal of Public Economics (170), 1–14.

Fevang, E., Hardoy, I., Red, K., 2017. Temporary disability and economic incentives. The Economic Journal l127 (603), 1410–1432.

French, E., Song, J., 2014. The effect of disability insurance receipt on labor supply. American Economic Journal: Economic Policy 6 (2), 291–337.

Fultz, E., 2015. Disability insurance in the Netherlands: a blueprint for U.S. reform? Washington, DC: Center for Budget and Policy Priorities.

García-Gómez, P., Gielen, A., 2018. Mortality effects of containing moral hazard: evidence from disability insurance reform. Health Economics 27 (3), 606–621.
Garcia-Mandicó, S., García-Gómez, P., O'Donnell, O., 2020. Earnings responses to disability insurance stringency. Labour Economics 66 (101880).
Gelber, A., Moore, T. J., Strand, A., 2017. The effect of disability insurance payments on beneficiaries' earnings. American Economic Journal: Economic Policy 9 (3), 229–61.
Gruber, J., 2000. Disability insurance benefits and labor supply. Journal of Political Economy 108 (6), 1162–1183.
Hausermann, S., Schwander, H., 2012. The Age of Dualization: The Changing Face of Inequality in Deindustrializing Societies. Oxford University Press, Ch. Varieties Of Dualization? Labor Market Segmentation and Insider-Outsider Divides Across Regimes.
Hullegie, P., Koning, P., 2018. How disability insurance reforms change the consequences of health shocks on income and employment. Journal of Health Economics 62, 134–146.
Hutchens, R., Grace-Martin, K., 2006. Employer willingness to permit phased retirement: Why are some more willing than others? Industrial and Labor Relations Review 59 (4), 525–546.
Karlström, A., Palme, M., Svensson, I., 2008. The employment effect of stricter rules for eligibility for di: Evidence from a natural experiment in sweden. Journal of Public Economics 92 (10-11), 2071–2082.
Koning, P., 2009. Experience rating and the inflow into disability insurance. De Economist 157 (3), 315–335.
Koning, P., Lindeboom, M., 2015. The rise and fall of disability insurance enrollment in the Netherlands. Journal of Economic Perspectives 29 (2), 151–172.
Koning, P., van Sonsbeek, J.-M., 2017. Making disability work? The effects of financial incentives on partially disabled workers. Labour Economics 47, 202–215.
Koning, P., van Vuuren, D. J., 2010. Disability insurance and unemployment insurance as substitute pathways. Applied Economics 42 (5), 575–588.
Kostøl, A. R., Mogstad, M., 2014. How financial incentives induce disability insurance recipients to return to work. American Economic Review 104 (2), 624–655.
Low, H., Pistaferri, L., 2015. Disability insurance and the dynamics of the incentive insurance trade-off. American Economic Review 105 (10), 2986–3029.
Maestas, N., Mullen, K. J., Strand, A., 2013. Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt. American Economic Review 103 (5), 1797–1829.
Marie, O., Vall Castello, J., 2012. Measuring the (income) effect of disability insurance generosity on labour market participation. Journal of Public Economics 96 (1-2), 198–210.
Moore, T. J., 2015. The employment effects of terminating disability benefits. Journal of Public Economics 124, 30–43.
Mullen, K. J., Staubli, S., 2016. Disability benefit generosity and labor force withdrawal. Journal of Public Economics 143, 49–63.
OECD, 2003. Transforming Disability into Ability: Policies to Promote Work and Income Security for Disabled People. OECD Publishing, Paris.
OECD, 2018. Public spending on incapacity.
URL https://www.oecd-ilibrary.org/content/data/f35b71ed-en
OECD, 2019. Society at a Glance 2019: OECD Social Indicators. OECD Publishing, Paris.
Ravesteijn, B., van Kippersluis, H., van Doorslaer, E., 2018. The wear and tear on health: What is the role of occupation? Health Economics 27 (2), 1057–9230.
Ruh, P., Staubli, S., 2019. Financial incentives and earnings of disability insurance recipients: evidence from a notch design. American Economic Journal: Economic Policy 11 (2), 269–300.

Salverda, W., 2020. De flexwolf verslindt de Nederlandse arbeidsmarkt. Economische Statistische Berichten 105 (4784), 178–181.

Staubli, S., 2011. The impact of stricter criteria for disability insurance on labor force participation. Journal of Public Economics 95 (9-10), 1223–1235.

Van Sonsbeek, J.-M., Gradus, R. H. J. M., 2006. A microsimulation analysis of the 2006 regime change in the Dutch disability scheme. Economic Modelling 23 (3), 427–456.

Van Sonsbeek, J.-M., Gradus, R. H. J. M., 2013. Estimating the effects of recent disability reforms in the Netherlands. Oxford Economic Papers 65 (4), 832–855.

Visser, M., Gesthuizen, M., Kraaykamp, G., Wolbers, M. H. J., 2018. Labor market vulnerability of older workers in the Netherlands and its impact on downward mobility and reduction of working hours. Work, Aging and Retirement 4 (3), 289–299.

Zaresani, A., 2018. Return-to-work policies and labor supply in disability insurance programs. AEA Papers and Proceedings 108, 272–276.

Appendix A Sensitivity analysis

Are the estimated reform effects sensitive to the cut-off number of days spent in the sickness scheme?

As explained in Section 3, due to the under-reporting of short-term sickness cases, we restricted the initial sample of sick individuals to include those who have spent at least 180 days in the sickness scheme to ensure that the three cohorts of sick individuals who fell sick in the third and fourth quarters of 2003 and the first quarter of 2004 are comparable in the number of days spent in sickness. The effects of the reforms could be under- or over-estimated if we allowed our study sample to include the sickness cases that last shorter than 180 days and are under-reported. For example, among the individuals with short-term sickness, those who are insured under the WIA may be inclined to resume working earlier than if they were insured under the WAO if they perceive that they eventually will not become entitled to DI due to a longer sickness period increasing the chances of recovery, but also due to much stricter eligibility criteria. In this case, the imposed sample restriction will drop proportionally more short-term sickness cases of relatively healthier individuals in the WIA group, and lead to an underestimation of the effects of the WIA reform.

Here we examine to which extent the estimated effects of the reforms are sensitive to allowing in the study sample short-term sickness cases that are shorter than 180 days and under-reported to different extents. The results from the regression given by Equation (1) on subsamples of individuals who have spent at least 90, 120, and 150 days in the sickness scheme show the following. With respect to labor participation and benefit receipt, when baseline results in Table 2a are compared to the sensitivity results in Table A.1a, we find small differences for the WIA reform in all outcomes except when the outcome is labor participation and the number of days spent in the sickness scheme is at least 90 days. Larger differences are observed for the transitional WAO reform when the outcome is labor participation, or when it is DI, and the number of days spent in the sickness scheme is at least 90 days. A potential reason is that, since under-reporting of short-term sickness cases occurs more often for the transitional WAO group compared to the WIA group, the estimates of the effects of the transitional WAO reform are more prone to be biased downward or upward, due to reasons as exemplified above. With respect to wage and benefit income, we observe similar differences for wage and DI income when baseline results in Table 2b are compared to the sensitivity results in Table A.1b. These results show that the baseline estimates of the effects of the WIA reform are to a large extent robust to the sampling of sick individuals with respect to the number of days spent in the sickness scheme. This is true to a lesser extent for the transitional WAO reform.

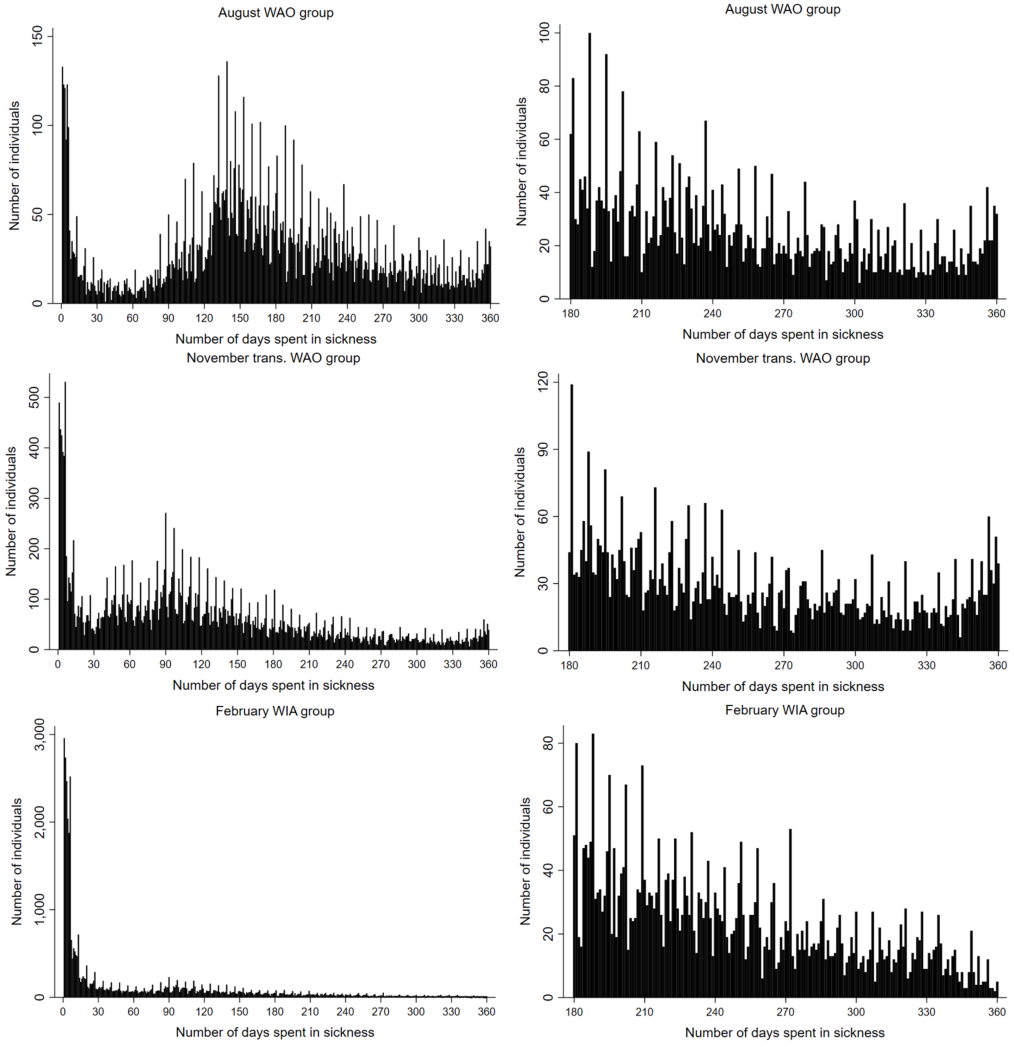

Figure A.1: Distributions of the number of days spent in sickness among individuals who fall sick in August 2003, November 2003 and February 2004, and participate in the WAO, transitional WAO, and the WIA, respectively. Left panel: distributions for the three groups when the number of days spent in sickness ranges from 0 to 360 where 360 is the maximum number of days individuals can spend in the sickness scheme if they participate in the WAO or the transitional WAO. Right panel: distributions when the number of days spent in sickness ranges from 180 to 360.

Table A.1a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt by the number of days individuals spend in the sickness scheme

		Number of days spent in the sickness scheme is at least 90	Number of days spent in the sickness scheme is at least 120	Number of days spent in the sickness scheme is at least 150
DI (and possibly UI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.021^{***}	-0.014^{***}	-0.010^{***}
		(0.002)	(0.002)	(0.003)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.052^{***}	-0.050^{***}	-0.053^{***}
		(0.002)	(0.002)	(0.003)
Labor participation	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.017^{***}	0.011^{***}	0.008^{**}
		(0.003)	(0.003)	(0.003)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.028^{***}	0.021^{***}	0.018^{***}
		(0.003)	(0.003)	(0.003)
UI (no DI) receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.002^{*}	0.002^{***}	0.003^{***}
		(0.001)	(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	0.010^{***}	0.011^{***}	0.013^{***}
		(0.001)	(0.001)	(0.001)
General assistance receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.001	-0.002	-0.002^{*}
		(0.001)	(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.003^{***}	-0.004^{***}	-0.004^{***}
		(0.001)	(0.001)	(0.001)
Other benefits receipt	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-0.000	0.000	0.001
		(0.001)	(0.001)	(0.001)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.004^{***}	-0.004^{***}	-0.004^{***}
		(0.001)	(0.001)	(0.001)
Observations		23,579,556	20,078,508	17,097,660
Individuals		125,535	106,807	90,918

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Table A.1b: Linear model explaining the effects of disability reforms on labor and benefit income by the number of days individuals spend in the sickness scheme

		Number of days spent in the sickness scheme is at least 90	Number of days spent in the sickness scheme is at least 120	Number of days spent in the sickness scheme is at least 150
DI (and possibly UI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-24.601^{***}	-14.913^{***}	-8.820^{**}
		(3.009)	(3.362)	(3.804)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-52.127^{***}	-46.730^{***}	-46.764^{***}
		(3.084)	(3.462)	(3.914)
Wages earned per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	41.900^{***}	23.891^{**}	16.379
		(10.100)	(10.895)	(11.799)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	77.722^{***}	59.118^{***}	53.001^{***}
		(10.528)	(11.407)	(12.339)
UI (excl. of DI) per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	3.820^{***}	4.767^{***}	6.152^{***}
		(1.447)	(1.574)	(1.721)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	16.603^{***}	18.725^{***}	21.714^{***}
		(1.489)	(1.627)	(1.783)
General assistance per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	-1.167)	-2.180^{*}	-2.847^{**}
		(1.147)	(1.245)	(1.367)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-4.193^{***}	-4.766^{***}	-4.952^{***}
		(1.187)	(1.297)	(1.426)
Other benefits per month	$\text{Treat}^{\text{Trans. WAO}} \times \text{Post}$	0.977	1.612^{*}	2.009^{*}
		(0.883)	(0.971)	(1.081)
	$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.567	0.095	0.145
		(0.888)	(0.983)	(1.088)
Observations		23,579,556	20,078,508	17,097,660
Individuals		125,535	106,807	90,918

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time and correspond to the months before individuals fall sick, and subsequently the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

How much are the estimated reform effects affected when the common trends assumption is violated?

In Section 5 we showed that in a number of outcomes the time trends of the control and treatment groups showed statistically significant differences at the end of the pre-reform period, from January 2002 to June 2003, providing evidence against the common trends assumption. Here we show that these differences in pre-reform time trends have no notable impact on the baseline estimates of the reform effects. We limit this analysis to the WIA group, and do not consider the transitional WAO group, since Section 5 showed that the common trends assumption is violated for this group for a larger number of outcomes.

We investigate the impact of the differences in the time trends of the control and treatment groups at the end of the pre-reform period on the baseline estimates of the reform effects as follows. We check whether the baseline estimates remain robust when we drop in our study sample the months where the pre-reform time trends of the treatment and control groups differ, but keep those months where they do not. To determine the months where the pre-reform time trends do not differ between the treatment and control groups, we first rely on visual evidence from Figures 2a and 2b. The figures suggest that the time-trends of the treatment and control groups overlap throughout the period from January 1999 to December 2001 in all outcome variables. To formally test this, we use regression analysis as in Section 5. That is, we use pre-reform data from January 1999 to December 2001, to estimate a regression where labor participation, benefit receipt, labor income, or benefit income is the outcome, and calendar month dummies, interactions of treatment and calendar month dummies, and time-invariant individual fixed effects are controls. Based on the F-statistic, coefficient estimates of interaction terms that are not statistically different from zero provide evidence in favor of the common trends assumption. The test results confirm that from January 1999 to December 2001 the pre-reform time trends of the treatment and control groups are statistically indistinguishable from each other for all outcomes except for general assistance income.[16]

Tables A.2a and A.2b present the estimates of the reform effects from the regression based on the restricted pre-reform period from January 1999 to December 2001. The estimates are very close to the baseline estimates in Tables 2a and 2b except that the estimate for general assistance receipt becomes insignificant, and the estimate for general assistance income becomes smaller and less significant. We conclude that the observed differences in the time trends of the control and treatment groups at the end of the pre-reform period have no notable impact on the baseline estimates of the effects of the WIA reform.

[16]The results for the F-statistics are as follows (p-values in parentheses). For labor participation, UI receipt, general assistance receipt, and other benefits receipt outcomes, the results, respectively, are 1.22 (0.180), 1.20 (0.193), 1.27 (0.135), 1.30 (0.114). For labor, UI, general assistance, and other benefits income outcomes, the results, respectively, are 1.31 (0.105), 1.05 (0.394), 1.48 (0.035), 0.89 (0.652).

Table A.2a: Linear model explaining the effects of disability reforms on labor participation and benefit receipt when the pre-reform period is restricted

	DI (and possibly UI) receipt	Labor part.	UI (no DI) receipt	General assistance receipt	Other benefits receipt
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-0.059^{***}	0.015^{***}	0.013^{***}	-0.002	-0.005^{***}
	(0.003)	(0.004)	(0.001)	(0.002)	(0.001)
Observations	12,916,711				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time. The pre-reform period corresponds to the months from January 1999 to December 2001. The post-reform period corresponds to the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Table A.2b: Linear model explaining the effects of disability reforms on the amounts of wage and benefits received per month when the pre-reform period is restricted

	DI (and possibly UI) received per month	Wages earned per month	UI (excl. of DI) received per month	General assistance received per month	Other benefits received per month
$\text{Treat}^{\text{WIA}} \times \text{Post}$	-50.829^{***}	55.893^{***}	20.480^{***}	-3.024^{*}	-0.000
	(4.414)	(13.836)	(1.894)	(1.671)	(1.239)
Observations	14,626,356				
Individuals	77,769				

Notes: All regressions employ the linear regression model with fixed effects given by Eq. (1) and include calendar month dummies for the periods of time before individuals fall sick and after they become eligible to apply for DI. January 1999 is the base month for comparison. The pre-reform and post-reform are periods of event time. The pre-reform period corresponds to the months from January 1999 to December 2001. The post-reform period corresponds to the months after individuals become eligible to apply for DI. Standard errors, in parentheses, are adjusted for clustering at the individual level. ***, **, * denote statistical significance at 1, 5, and 10 percent, respectively.

Are the estimated long-term effects of the reforms sensitive to the correlation between the event and calendar time dummies?

In Section 5 we analyzed the long-term effects of the disability reforms using the regression given by Equation (2). In particular, the long-term effects are analyzed through the interactions of treatment and year dummies over a period of ten years. In this equation, the year dummies of event time are correlated with the month dummies of calendar time by construction. This correlation structure could affect the point estimates, or the standard errors, of the main effects of interest given by the coefficients of the interaction terms. To check whether the correlation between the year dummies of event time and month dummies of calendar time lead to biased estimates of long-erm effects, instead of year dummies, we consider dummies that indicate periods of three years of event time which have much weaker correlations with the month dummies of calendar time. Figures A.2a and A.2b present the results based on three-year period dummies. Using these dummies leads to estimates of the effects of the interaction terms very similar to those when year dummies are used in Figures 6a and 6b so that our qualitative results are not affected.

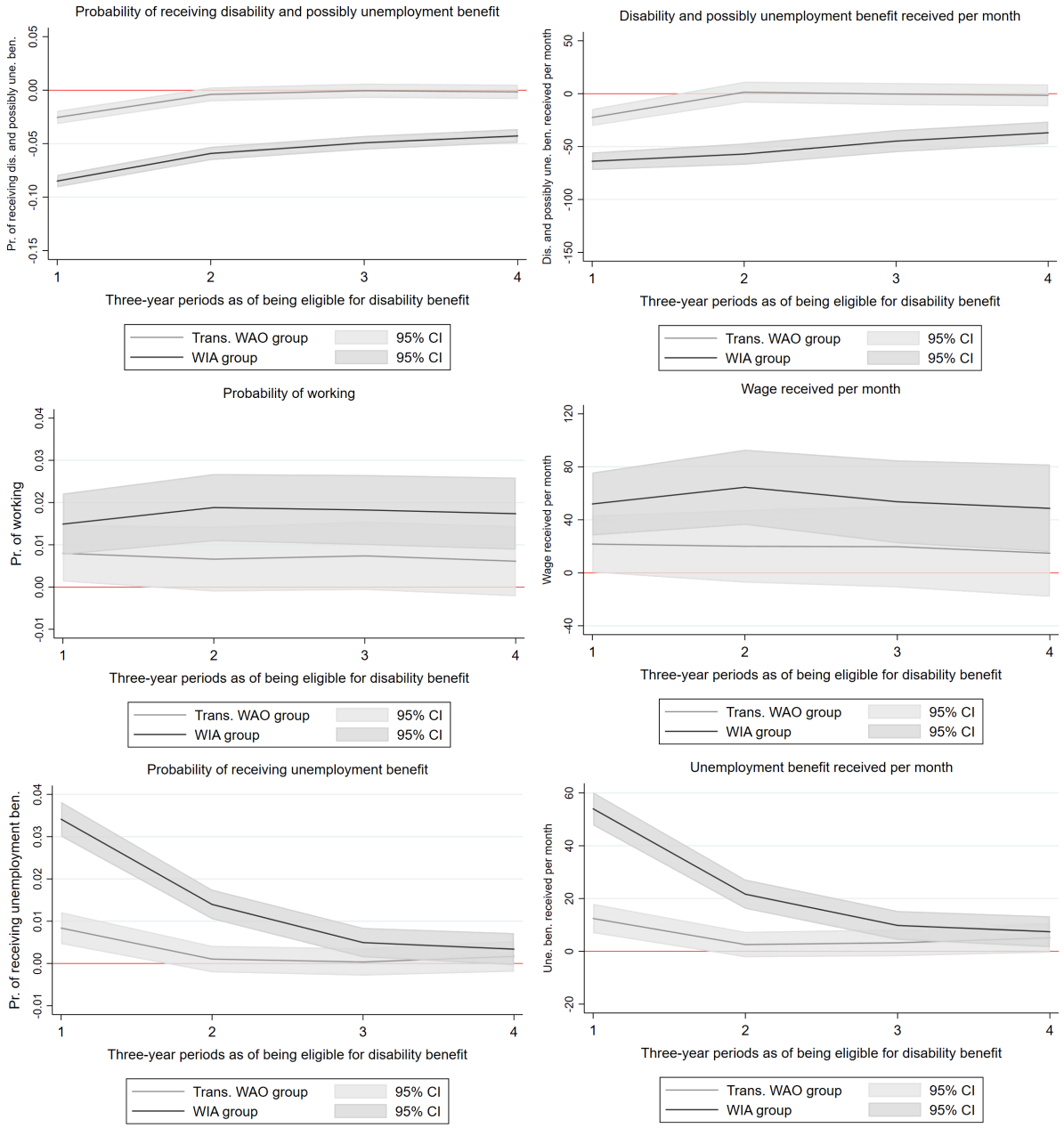

Figure A.2a: Coefficient estimates of treatment and annual dummy interactions for control and treatment groups during the reform period from regressions for labor participation, benefit (DI and UI) receipt, labor income, and benefit income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Labor participation, benefit receipt, labor income or benefit income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.

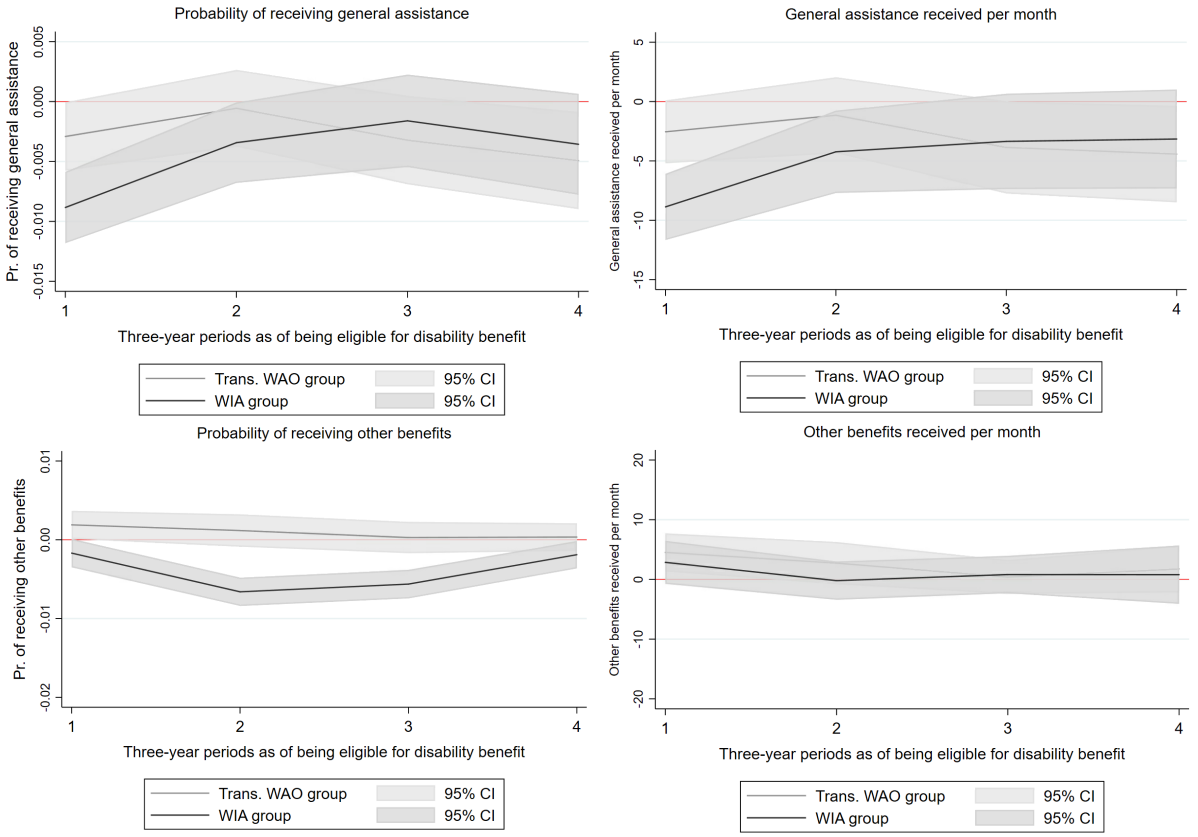

Figure A.2b: Coefficient estimates of treatment and annual dummy interactions for control and treatment groups during the reform period from regressions for benefit (general assistance and other benefits) receipt and income. Around each estimate is a 95 percent confidence interval. Regressions are based on the data available on a monthly basis for the reform period of ten years. The pre-reform period is the base for comparison for annual dummies (Section 5). Benefit receipt or income is the outcome, and annual dummies, treatment and annual dummy interactions, and time-invariant individual fixed effects are controls. Each regression uses 14,626,356 observations for 77,769 individuals who fell sick during the period from July 2003 until March 2004. Standard errors are adjusted for clustering at the individual level.